Journal of
Experimental Psychology

VOL. 28, NO. 2 FEBRUARY, 1941

RETROACTIVE INHIBITION AS A FUNCTION OF DEGREE OF GENERALIZATION BETWEEN TASKS[1,2]

BY ELEANOR J. GIBSON

Smith College

A hypothesis relating verbal learning to two concepts from the literature of conditioning—generalization and differentiation—has recently been presented by the writer (4). This hypothesis has advanced the general notion that *discrimination* is an essential process in verbal learning, and that the special difficulties referred to by such terms as *interference* and *retroactive inhibition* can be understood, at least in part, as cases of relatively low discriminability of learning material. When this notion is combined with the principle of

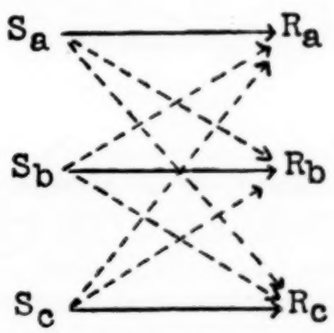


FIG. 1

establishment of differentiation between the previously undiscriminated items, many implications follow. These implications are stated as specifically as possible in the form of predictions in the article referred to. Figure 1 below may clarify the mechanism assumed to be at work. It represents a list of paired associates in

[1] This experiment is one of a series of studies presented to the Faculty of the Graduate School of Yale University in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The writer wishes to acknowledge the valuable advice and assistance of Professor Clark L. Hull during the progress of this work.

[2] This experiment was reported at the 1938 meeting of the American Psychological Association (3).

which generalization is assumed to be taking place during learning. Generalization is defined as "the tendency for a response R_a learned to S_a to occur when S_b (with which it has not been previously associated) is presented" (4, p. 204). In other words, S_a, S_b and S_c possess low discriminability when they are presented separately. Generalization tendencies (represented by broken arrows) will tend to block the right responses (solid arrows) in proportion to the strength of the tendencies. Competitive blocking thus occurs, rendering learning difficult.

In a two-list situation also, generalization and competitive blocking may occur. Figure 2 represents the relationships. When the second list is learned, its stimulus members may generalize with those of the first list, and the responses of the first list tend to occur, or to block the right responses, during the new learning. Likewise, if List 1 is recalled after List 2, generalized responses from List 2 may interfere with recall, producing competitive blocking or overt generalization (reversion to the wrong list). The right excitatory

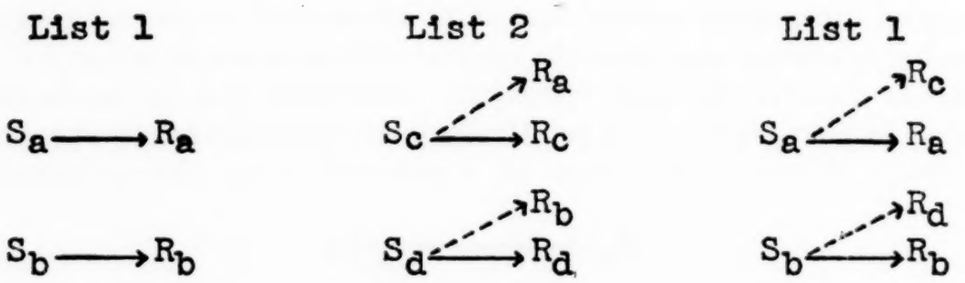


FIG. 2

tendencies will be weakened, it is assumed, in proportion to the degree of inter-list generalization. This argument leads to the prediction in the hypothesis (4, p. 215) that "a first list will be more poorly recalled as the strength of the tendency for items of a second list to generalize with it increases." The present experiment is designed to test this proposition; that retroactive inhibition is a function of degree of generalization between a primary and an interpolated list. The experimental situation will yield data relevant to the prediction of interference from List 1 to the learning of List 2 as well as retroactive inhibition. The hypothesis predicts that "more repetitions will be required to learn a second list, in proportion to the strength of the tendency for items of a first list to generalize with items of a second list."

Numerous experiments have demonstrated that retroactive inhibition varies with the degree of similarity between an original and an interpolated task (1, p. 389 ff.). The present hypothesis is obviously

related to this work, but would maintain that *indistinguishability*, rather than *similarity* or *homogeneity* (14) is the important variable. Similarity implies a sophisticated phenomenal relationship involving both the likeness and the discriminability of the items in question. But it is the indistinguishability of the items, and the necessity for differentiating them which forms the basis of the present hypothesis. At the outset of learning, items usually have varying degrees of stimulus equivalence, particularly in the verbal learning situation where items are not presented side by side for perceptual comparison. It is reasonable to suppose that this equivalence or indistinguishability is essentially similar to the classical fact of generalization which appears in the early stages of conditioning.

Since the purpose of the present investigation was to discover the relationship between generalization and retroactive inhibition, a preliminary experiment was designed to standardize material in terms of degree of generalization. The method of standardization and the resulting data are presented in Experiment I. The standardized material is utilized in Experiment II, a retroactive inhibition experiment, to form series of primary and interpolated lists differing in degree of generalization. Experiment III is a repetition of Experiment II using individual subjects instead of the groups used in II.

Experiment I. Determination of Degree of Generalization of Stimulus Items

Procedure

For determining degree of generalization of items, a technique previously employed by Yum (15) was used. A series of *standard* forms was selected, to be used as cues in paired associates learning. A number of variations from each of these forms was then drawn. Twenty-four hours after a group of subjects had learned the list of paired associates, a recall series was given them, in which each cue form was either the original standard, or a *variation* substituted for the standard. The measure of generalization was taken as the extent to which the subjects responded to the variation *as if it were the standard.*[3] Yum was interested in relating the frequency of response to the variation with the degree of perceptual similarity between the variation and the standard. Perceptual similarity was determined by judges' ratings. To make the present experiment comparable, a number of groups of forms, each group consisting of a standard and several variations, were drawn on cards and presented to ten judges. They were asked to put each variation with the standard which it resembled most. When this was finished, they were asked to rank the variations according to degree of similarity to the standard. Thirteen groups of forms were treated in this way. Each finally consisted of a standard form and two variations from it. There were available then three classes of forms: standard, first degree similarity, and second degree similarity. Included also in the experiment was a fourth class of forms which bore no resemblance to any of the standards, as determined by the judges.

Relative degrees of generalization for the variation-forms were now determined. The 13 standard forms were used as the cue members of a list of paired associates. Each standard was

[3] This criterion is appropriate to our definition of generalization—"the tendency for a response R_a learned to S_a to occur when S_b (with which it has not been previously associated) is presented" (4, p. 204).

paired with a monosyllabic word (again in accordance with Yum's procedure). The order of the pairs was varied with each presentation, so that no serial associations would be formed. Four *test* lists for determining degree of generalization were prepared, each composed of 13 forms, like the learning list, but forms of all 4 classes were included among them. In a typical test list there would be 3 of the standard forms, 3 forms highly similar to their standards, 3 forms less similar to their standards, and 3 forms which resembled none of the standards. (Actually, there had in any list to be 4 forms from one of the classes to make up 13.) There were necessarily 4 different test lists, to include the total of 54 forms. No variation ever appeared in the same list with its standard. Four groups of subjects took part in the experiment, varying in number from 19 to 29. Each group had the original learning list, but a different test list. The groups were four sections in an elementary psychology course.

A special serial exposure device was used to present the material to the subjects, consisting of a Jastrow memory apparatus placed behind a projector which threw a large image of each form-word pair on a translucent screen. The forms had been printed by mimeograph so that perfect copies were available for the different lists, and the response words were printed in black India ink. Exposure time for each pair was 2 seconds. Five learning trials of the original list were given; immediately following each presentation, recall was tested, the forms alone being presented in a different order each time. The subjects had to record the proper response words in a prepared booklet within 3 seconds. The subjects were then released but returned the next day at the same hour and were given one of the four test lists for determining degrees of generalization. This test was similar in procedure to one of the previous recall trials. Following this test series, an ordinary recall test was given using the original forms.

The following instructions were given the subjects at the beginning of the experiment:

> You will be shown a group of forms, each one paired with a word. Study these pairs so that when a form is shown alone, you can write the appropriate word. You will be shown only one pair at a time. Do not try to learn these pairs in any particular order, because the order will be changed every time. The point is, to *associate* a particular word with the form with which it always appears.
>
> After every presentation of the forms paired with words, you will be shown the forms by themselves in order to see whether you remember the appropriate word or not. If you do remember it, write it in the blank provided, being sure to put it in the proper space. Do not leave any blank spaces; if you do not remember the proper word, draw a line through the space. After we have finished one of these test trials, turn the page so that you always have a blank page before you during the learning.

On the second day, before the test series, these instructions were given:

> Today I want you to look at some forms as they appear and put down the word for each one whenever you can. You will be shown forms by themselves; try to write the appropriate words, just as you did in the test series yesterday. You may not remember them all the first time, but you will be given another chance later.

The subjects could hardly fail to recognize at some point in the series that this was not the original list, since 3 or 4 of the forms were totally unfamiliar ones bearing no resemblance to the standards. This awareness probably had the effect (to judge from later questioning) of causing them to write responses only to those forms which they felt 'sure' of, so that in general when variations were responded to as if to the standards, the items were functioning as *equivalent* stimuli, not as merely recognizably similar stimuli.

Results

Degree of Original Learning.—Since variations in the degree of learning attained are probably accompanied by variations in the degree of generalization exhibited,[4] the mean number of response words correctly recalled after the fifth learning trial for each group

[4] Hilgard and Marquis (6, p. 182) present evidence to show that with conditioned responses the curve of generalized responses differs with different stages of practice.

of subjects is presented in Table I. The degree of learning attained

TABLE I

MEAN NUMBER OF RESPONSE WORDS CORRECTLY RECALLED AFTER 5 PRESENTATIONS

Group I	Group II	Group III	Group IV
12.5	12.0	11.9	11.9

was high, being approximately 12 out of a possible 13 correct responses. This implies that many individuals were able to recall all 13 responses correctly. The differences between the averages of the 4 groups are satisfactorily small.

Degrees of Generalization for Individual Forms.—The main purpose of the results was to establish four lists of forms varying in their tendency to call forth the same responses as did the original list of standard forms. To obtain such lists, the results were scored as follows. First, the number of subjects responding correctly to each standard form in the test list was obtained. Then, the number responding to each of the *variations* as if it were the standard was calculated. When these latter figures are expressed as percentages of the total number taking part in the group concerned, they constitute measures of the degree of generalization with the standard. Finally, it was determined whether any form rated dissimilar to the standards was responded to as if it were one of the standards; and whether any variation from a standard was responded to as if it were a different standard.

Figure 3 shows the forms grouped in classes according to their tendency to generalize with the standards. The standard forms themselves are Class I. The forms in Class II are those variations which produced the highest frequencies of responses belonging to their standards. Those grouped in Class III produced lower frequencies of the standard responses. In each case, the forms in Class II and Class III are listed next to their standards. The percentage of the group which responded appropriately is in each case given opposite the form. The means show a wide difference between Classes I, II, and III in terms of tendency to call forth the standard response. The forms grouped as Class IV are those which were judged dissimilar. Four of these forms were responded to by a few subjects, but with no consistency as regards the responses given. If the mean percent response is calculated regardless of the nature of the response, the figure is 2.8 percent. The order in which forms of this class are listed is necessarily arbitrary. None of the variations (Classes II and III) was responded to with any consistency as if it were a standard other than the one beside which it is listed in the chart. There were a total of 10 'wrong' responses each for Class II

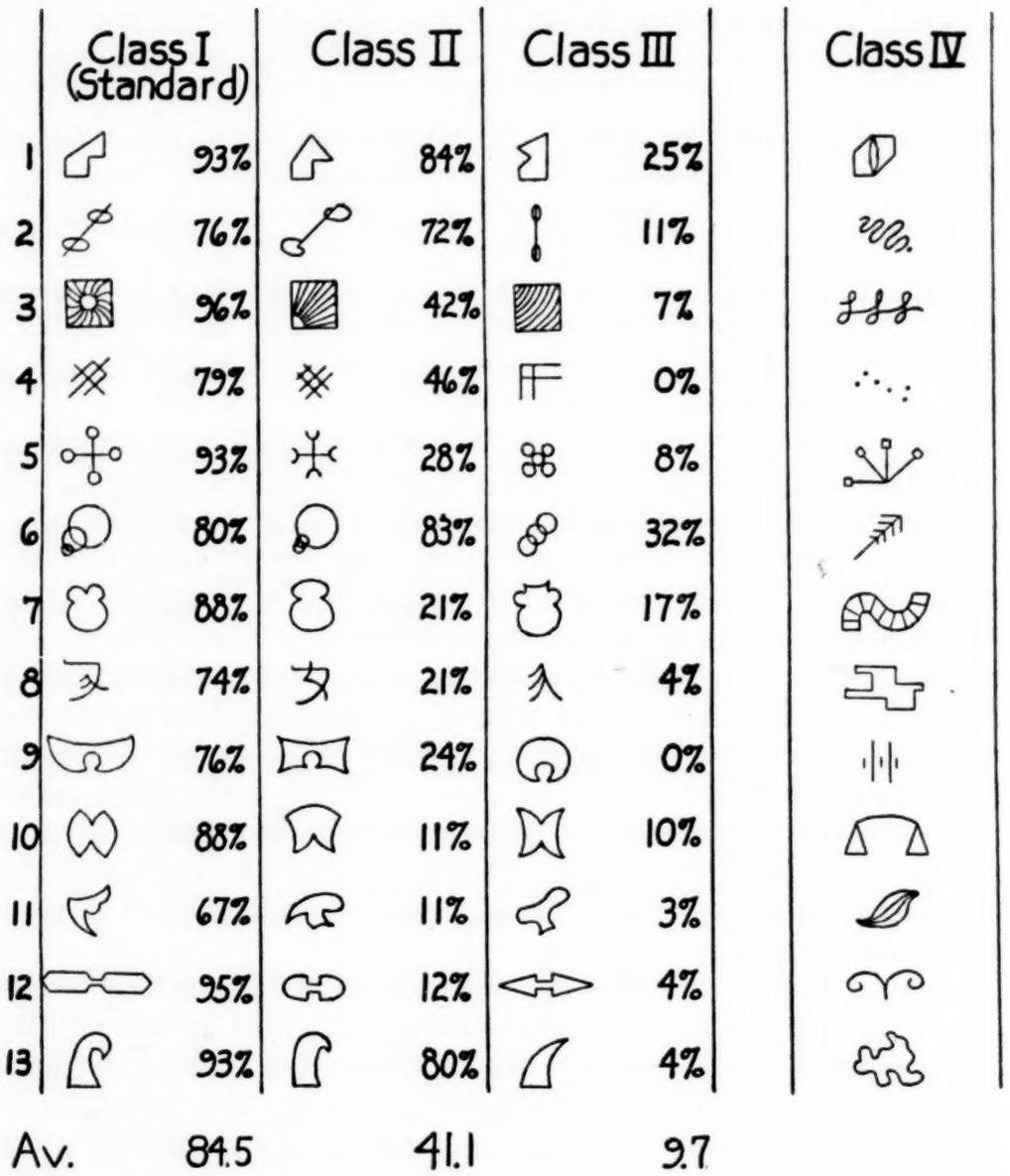


FIG. 3. Stimulus forms grouped in classes according to the percentage of subjects giving the standard responses.

and Class III—about 10 percent of the total possible responses in each case.

Two of the forms included in Class III were not responded to at all, although these two forms had been judged similar to the standards next which they are listed. Besides these two cases, there were only two discrepancies between similarity as rated by the judges and generalization as determined by quantitative score. These discrepancies caused a change from Class III to Class II in the case of forms 10 and 12. There was therefore a high correlation between judgments of similarity and objective tendency to generalize. To show this correlation, the results can also be scored according to a

method used by Yum. He calculated for each test list the mean number of forms of each *similarity* group responded to 'correctly,' rather than scoring each form individually as has been done above. If this calculation is made in the present experiment, the data are consistent for the four test lists in showing the highest mean response to the standard forms, next highest to forms judged of first degree similarity, and least high to those judged of second degree similarity.[5]

Experiment II. Degree of Retroactive Inhibition and of Interference as Dependent on Degree of Generalization

Procedure

Since it was proposed to test the relationship between degree of inter-list generalization on the one hand, and degree of retroactive inhibition on the other, the conditions of the experiment were planned so as to include three different degrees of generalization between primary and interpolated lists (Conditions II, III and IV below). Another condition contained identical stimulus items in the two lists (Condition I), and a fifth served as a control for calculating retroactive inhibition (Condition V). In all conditions, the primary learning list was made up of the standard (Class I) forms paired with nonsense syllables. The four classes of forms as they are grouped in Fig. 3 constituted the stimulus items for the four interpolated lists. The plan of the experiment was then as follows. The response members of the interpolated lists were nonsense

Condition	Learning of Primary List	Learning of Interpolated List	Retention of Primary List
	Stimulus Forms	Stimulus Forms	Stimulus Forms
I	Standard	Class I (Standard)	Standard
II	Standard	Class II	Standard
III	Standard	Class III	Standard
IV	Standard	Class IV	Standard
V	Standard	(No learning)	Standard

syllables different from those of the primary list, but identical in all four interpolated lists. Conditions I, II, III, and IV differed, therefore, only as to stimulus forms. The syllables for both primary and interpolated lists were chosen according to the usual rules from Hull's list (7). Only those of low association value were included. In condition V, the subjects were given copies of the '*New Yorker*' to read during the interpolated interval, with instructions to select the best and the worst cartoon.

The details of procedure were similar to those of Experiment I. Since the nature of the five conditions precluded using a subject more than once, five large, equivalent groups of subjects were secured. The sections of an elementary laboratory course offered the most available and suitable subjects. The same technique of exposure was employed, again allowing two seconds per pair for learning, and three seconds per item for the recall trial which followed. The order was varied with every learning trial and every test trial. Five learning and test trials were given.[6] The

[5] The means for the four lists combined are: Standard forms, 2.47; first degree similarity, 1.12; second degree similarity, .33. The maximum response would be 3, since each test list contained forms from each of the four groups. The data are presented in detail in (2).

[6] Five repetitions were chosen on the basis of a preliminary experiment. It allowed only partial learning of the lists (about 8 out of 13 items) which was desirable in order to secure a high degree of retroactive inhibition (8).

prepared booklets in which responses were recorded were collected immediately following the original learning, and new booklets passed out. Instructions for the interpolated learning followed and this task was begun at once. The interpolated list was also presented five times, with tests following each presentation. As soon as the interpolated learning was finished, books were collected, others distributed, further instructions given, and recall and relearning of the original list was begun. In all five conditions, the interpolated period between the end of the fifth test of original learning and the first recall test of it lasted fourteen minutes. Seven relearning repetitions of the first list were given. The number was limited to seven, owing to the termination of the class period; it did not allow learning to complete mastery in most cases.

The instructions were almost identical with those given in Experiment I and need not be repeated. After the 5 presentations of the first list, subjects were told: "We shall return to this list later on, and give enough more trials for everyone to finish learning." When the new record books had been handed out, they were told: "Now there will be a new list, of the same type, to be learned. The procedure will be exactly similar to that we have just followed. Please try again to work as hard as possible." The final instructions given before recall of the first list were these: "Now we shall go back to the first list and repeat it until everyone has learned all of the pairs. We shall start with a test trial to see how much you remember. Write every syllable which you can recall in the proper space when its form is shown, and draw a line in the others, as usual. Then we shall proceed with alternate learning and test trials as we did before."

Results

Degree of Original Learning.—Since the validity of the results depends on the equivalence of the 5 groups of subjects as regards degree of original learning, the mean number of syllables correctly recalled on the final test for each condition is presented in Table II.

TABLE II

DEGREE OF ORIGINAL LEARNING

Condition	I	II	III	IV	V
Mean No. Learned.....	8.21	7.86	8.05	7.92	7.91
No. of *S*'s.....	24	22	22	39	23

The means are fairly uniform, and the variation could not be responsible for the differences in recall which will appear later.

Degree of Learning of Interpolated Lists.—The degree of interpolated learning achieved in conditions I, II, III, and IV is interesting on two scores. First, the hypothesis predicts that higher degrees of generalization between a first and a second list should make the second list harder to learn. And second, variation in the degree of interpolated learning achieved may in itself affect the degree of retention of the primary list, since degree of retroactive inhibition has been shown to be a function of degree of interpolated learning (**9, 12**). Table III shows the mean number of syllables correctly recalled after the fifth trial of the interpolated list in each of the four conditions, as well as the mean total number recalled on all five tests of interpolated learning. Only the differences between Condition I

TABLE III

DEGREE OF LEARNING AND MEAN NUMBER OF REVERSION ERRORS IN INTERPOLATED LEARNING

Condition	I	II	III	IV
Mean No. Recalled at Trial 5.....	7.67	9.68	10.00	9.82
Mean Total No. Recalled.....	22.75	30.91	31.82	33.95
Mean No. Reversion Errors.....	.83	.45	.41	.41

and the other conditions approach statistical reliability,[7] but the totals show a definite tendency for more syllables to be correctly recalled as degree of generalization decreases from Condition I to Condition IV. The learning curves for each of the four conditions are presented in Fig. 4. Condition I is lower than the others throughout the five trials. In the early trials, Condition IV is noticeably higher than the other three, but in the later trials its curve is close to the curves for Conditions II and III. The chief fact shown is the much lower degree of learning achieved in Condition I, which is in keeping with the hypothesis under consideration. It might also affect the degree of retroactive inhibition for this condition by lowering it, since it has been shown that retroactive inhibition decreases with lower degrees of interpolated learning (within the range of partial learning).

When the responses given during interpolated learning are analyzed for reversion errors (cases of overt generalization with the preceding list) it should be expected, according to the hypothesis, that more such errors accompany the condition where learning is most difficult. The mean number of such errors for the four conditions are also given in Table III. It is obvious that more such errors occur in Condition I, coinciding with the greater difficulty of learning in that series. The greatest difference between conditions according to both recall and reversion measures, comes between Condition I and Condition II. A striking feature of the reversion errors, however, is the low frequency with which such errors occur at all, overtly.

Degree of Retroactive Inhibition.—The degree of retention of the first list may be calculated from several measures: from the first recall test immediately following interpolated learning (Recall 1), from the second recall test, which follows the first relearning trial (Recall 2), and from a relearning score. Table IV gives the percent of retroactive inhibition based on the number of correct recalls at the first and second test.[8] All four experimental groups show a

[7] The critical ratio ($D/\sigma_{diff.}$) of the difference between Conditions I and IV is 2.84 when degree of learning is measured by the number recalled at trial 5, and 3.75 when measured by the total number recalled.

[8] The retroactive inhibition scores have been calculated by the usual formula from the percent retained, rather than directly from the number recalled.

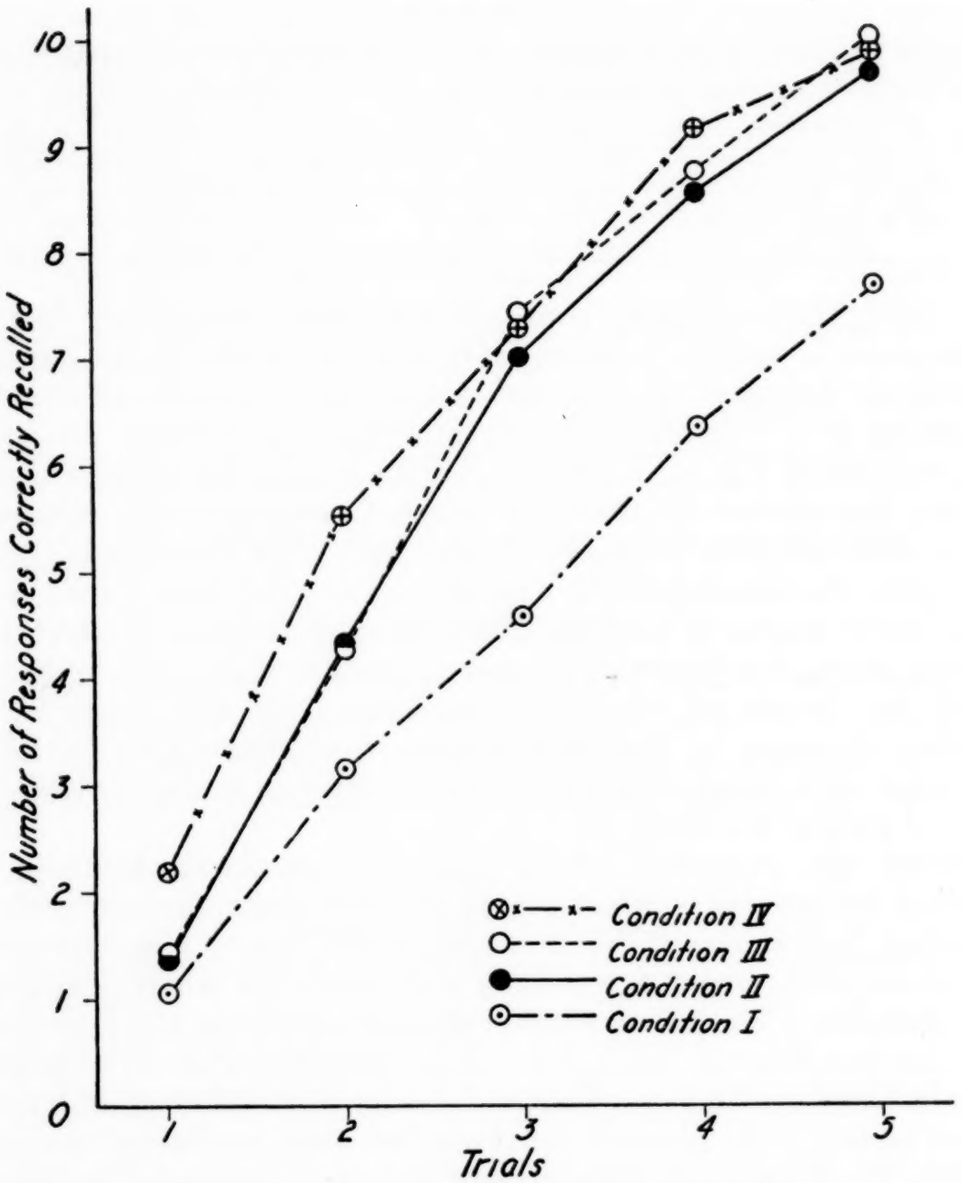


FIG. 4. Learning curves of the four interpolated lists. Degree of learning is plotted in terms of the number of responses correctly recalled after each trial.

TABLE IV

RECALL SCORES, PERCENT OF RETROACTIVE INHIBITION MEASURED BY RECALL SCORES, AND MEAN NUMBER OF TRIALS TO RELEARN 9 ITEMS

Conditions	I	II	III	IV	V
Mean of Recall 1	3.04	3.00	4.18	5.79	7.57
Mean of Recall 2	5.92	6.55	7.00	8.36	9.78
Retroactive Inhibition, Recall 1	60.6%	59.9%	45.8%	23.5%	
Retroactive Inhibition, Recall 2	41.7%	32.6%	29.6%	14.6%	
Mean Trials to 'Relearn'	4.08	3.73	3.86	3.00	2.35

significant degree of retroactive inhibition in terms of both measures of recall. Furthermore, the degree of retention varies with the degree of generalization between the original and the interpolated learning.[9] Figure 5 shows degree of retroactive inhibition plotted against the

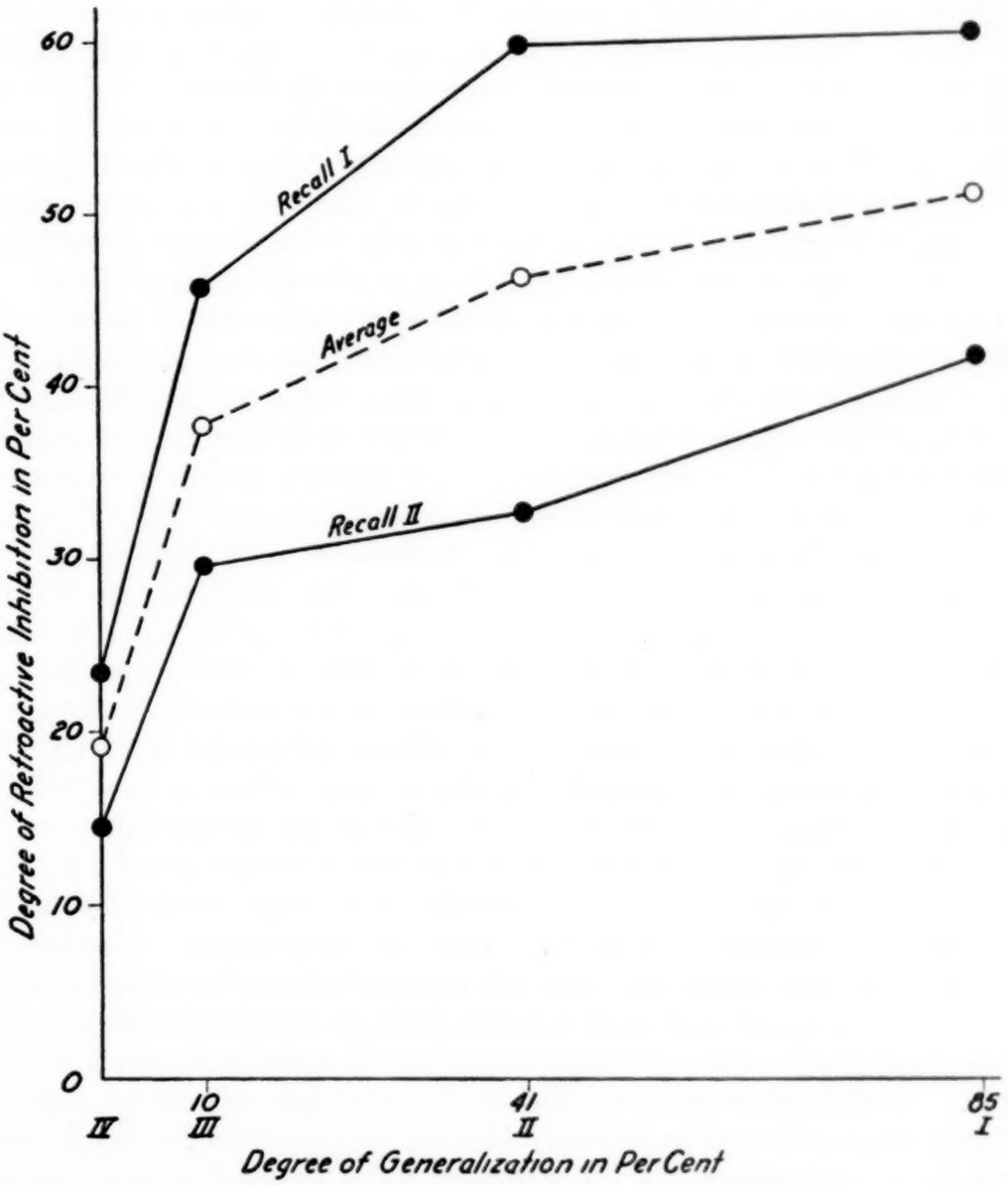


FIG. 5. Degree of retroactive inhibition plotted as a function of degree of generalization between stimulus forms of original and interpolated list. The Roman numerals under the percentages of generalization on the abscissa indicate the experimental condition previously described.

[9] Critical ratios ($D/\sigma_{diff.}$) based on Recall 1 are: I and III, 1.69; I and IV, 3.83; I and V, 5.56 II and III, 1.56; II and IV, 3.51; II and V, 5.17; III and IV, 2.11; III and V, 3.96. These statistical reliabilities are in general satisfactory, considering that a gradient is involved. Critical ratios for Recall 2 range from 1.18 to 3.07 when alternate conditions are compared.

degree of generalization as obtained in the standardization experiment. As the degree of generalization increases, so also does degree of retroactive inhibition, as the hypothesis predicts. When the two curves are averaged, the result is a smooth, negatively accelerated gradient.

It is noteworthy that Condition I (identical stimulus items) does not show a significant increase in retroactive inhibition over Condition II (Class II stimulus items), particularly in Recall 1. A greater difference is apparent in Recall 2, but even here the critical ratio is only .74. This result may be due to the lower degree of interpolated learning for Condition I. In other words, had degree of interpolated learning in Condition I been as high as that in the other three series, a higher degree of retroactive inhibition might have been found.

Another feature of the curves is the sharp rise from Condition IV to Condition III—from 0 generalization to 10 percent generalization. Yet Condition IV itself shows 23.5 percent retroactive inhibition in the first recall. It may be questioned whether such a degree of retroaction for a condition of supposedly zero generalization is consistent with the theoretical prediction. Two explanations suggest themselves. The first is that syllable responses were learned in the interpolated list as well as in the primary list, and that something analogous to generalization may take place among *response members* as well as stimulus members. Doubt is cast on this possibility by the McGeochs' study (**10**), which showed that variation in degree of similarity of response members of interpolated paired associates to response members of a primary list was without effect in determining degree of retroactive inhibition. This fact is consistent with Thorndike's finding (**13**) that similar response members are not frequently substituted for one another in recall, but that similar stimulus members do frequently evoke one another's responses. The second possibility is that there was actually generalization between stimulus members of original and interpolated lists in Condition IV, in spite of the negligible tendency established in the standardization experiment. It will be remembered that a very high degree of original learning was achieved in the standardization experiment, while only 8 out of 13 responses, on the average, were learned in the present experiment. It is suggested that generalization, after reaching a maximum, decreases as learning increases; and that as generalization decreases, the stimuli most discriminable from the standard are the first to cease producing generalized responses. In this case, the forms belonging to Class IV were ineffective in producing generalization in the standardization experiment because the degree of learning was high and because they varied considerably from the standards. But in the present experiment, they were effective to some extent

because the degree of learning was much lower. Evidence to support this explanation may be found in the fact that some cases of errors of reversion to the interpolated list actually did occur in Condition IV, though not as many as in the other conditions. But before this explanation can be accepted, it will be necessary to study the curve of generalization in relation to degree of learning under comparable conditions.[10]

The tendency for a greater decrement to accompany higher degrees of generalization is further supported by the number of trials required to relearn to a set criterion. This measure is of doubtful value, since learning was not carried to a criterion originally. But it should be of some significance, since the groups were equated as to number learned, on the average. Table IV gives the mean number of trials required in each condition to relearn to a criterion of 9 out of the 13 items. (A higher criterion could not be set, since some subjects learned no more than 9 in the seven relearning trials.) Condition I required more trials, on the average, to relearn 9 items, than did any of the others. There is not a perfect gradient here, since the average for Condition III is slightly higher than that for Condition II. It is interesting to note that the difference between Condition III and Condition IV is again the greatest, as was the case with the two other measures of retention.

Errors in Recall.—In the light of the above results, it is interesting to examine the mean number per subject of errors of reversion to the interpolated list which occurred in recalling the first list in the four conditions. The figures for the four conditions are: Condition I, 1.00; Condition II, 1.09; Condition III, .91; Condition IV, .36. As was the case when retention was measured by first recall, there is little difference between Condition I and Condition II. What is especially interesting is the large difference between Condition III and Condition IV, which coincides with the great difference between these conditions in all three measures of retention.

It is obvious that actual reversions account for only a small percentage of the errors which occur in either interpolated learning or recall and relearning of the primary list. Superficially, this fact may appear to be out of line with the hypothesis under test. But when the nature of the other errors is studied, their occurrence seems compatible with the hypothesis. Omissions are the most frequent type of error. When the relearning trials in Condition I are examined, omissions account for 67.9 percent of the errors. Theoretically, omissions may occur because there is *no* strong excitatory tendency to respond, or because the tendency or tendencies are

[10] This problem has been taken up in an investigation to be published.

inhibited. The former explanation may account for a number of the omissions, since the primary list was not learned to a very high degree originally. The second explanation is in line with the hypothesis, since it would assume that when a correct response and a generalized response, or two generalized responses both tend to occur at once, blocking may result. It is also possible that a response, although learned, was extinguished during the interpolated learning and is thereby inhibited. Such a possibility requires an extension of the hypothesis which will be considered later.

The errors other than omissions may be divided into three types—reversions or cases of overt generalization *between* lists, cases of overt generalization *within* the list, and misspellings. During the relearning of Condition I, 3.2 percent of the errors were reversions; 5.6 percent were cases of overt generalization within the list; and 23.4 percent were misspellings of the syllable responses. The question arises as to whether the errors of misspelling are consistent with the hypothesis. When these errors are examined carefully, a very interesting discovery emerges. Many of them appear to be compromises resulting from conflicting excitatory tendencies. For instance, two forms within a list have as their responses *ruv* and *tah*, respectively; the subject, instead of responding correctly to either of them, may respond with *tuv* to both. Or, a form in the interpolated list which has been shown previously to generalize with a form in the primary list, may have as its response *kiv*; the form in the primary list may be paired with *haj*. But when the subject is asked to recall the primary list, he may write neither of these responses, but instead *kij*. In such cases, the stimulus form seems to have two excitatory tendencies, an associated one and a generalized one, which result in a compromise response rather than blocking. Under what conditions compromise will occur instead of blocking is an intriguing question. It seems possible that the strength of both tendencies may be the key to this problem—that two equal but weak tendencies may lead to compromise, strong ones to blocking—but further study is required to solve it. Probably not all the misspelling errors are of this type; an attempt was made to estimate the percentage falling in this class, but it proved impossible to classify many of the cases with certainty.

The errors of generalization *within the list* which occurred during the learning of the primary and the interpolated lists show an interesting trend. There is a mean decrease of 59 percent in such cases from primary learning to interpolated learning. This decrease occurs consistently in all the four conditions. The reduction in Condition I is consistent with a prediction regarding the learning of pre-differentiated items—"If differentiation has been set up within a list, less generalization will occur in learning a new list which includes the

same stimulus items paired with different responses" (4, p. 222). The reduction in the other conditions is hard to understand unless something like transfer of differentiation is postulated.

Experiment III. A Repetition of Experiment II Using Individual Subjects

Procedure

Since the third experiment was undertaken principally as a check on the reliability of the results obtained in Experiment II, conditions were kept as nearly the same as possible. In general, only those changes were made which were demanded by individual experimentation in place of a group procedure. The two major changes were omission of Condition V, and the use of a learning criterion rather than a set number of trials. Condition V (the control for retroactive inhibition in which the subjects read during the interpolated period) was omitted, because there could be no doubt of the fact that retroactive inhibition occurred and we wished only to verify the gradient observed from Condition I to Condition IV. The subjects always learned to a criterion in the primary learning, the interpolated learning, and the relearning so as to insure equal degrees of learning in the four conditions.

The lists learned by the subjects in Conditions I, II, III, and IV were identical with those of Experiment II, except that one pair was omitted, making them 12-unit instead of 13-unit lists. The forms were pasted on heavy paper, and the syllables were typed beside them. The learning material was presented on an electrically driven Chicago-type memory drum. Each pair was again exposed for two seconds. But when the forms were presented alone, in the test trials, they now appeared for 2 seconds each instead of 3 seconds, since the subjects had only to speak the right associates instead of to write them. Six random orders were prepared for each list, and the order was changed each time for both learning and test trials. All the subject's responses were recorded by the experimenter.

The procedure was as follows. As soon as the subject had received the instructions, which were the same as those of Experiment II, the first list was presented. When the list had been exposed, the forms alone were presented and the subject attempted to recall the appropriate responses. This same procedure was continued until the subject had recalled 8 of the 12 responses correctly in a single trial. Then the subject was given a four minute rest period while the list was changed. In order to prevent rehearsing, he was given a magazine to read. At the end of this period, the interpolated learning was begun. The procedure was exactly the same as during original learning, and was continued until the subject had learned 8 of the 12 responses. Another rest period ensued, and the subject read, as before, until the time interval was up. The total time allowed for the interpolated interval was 20 minutes. The subject then tried to recall the responses of the first list, and relearning followed. Relearning was carried to a criterion of one perfect recall trial. The experiment required about one hour.

Fourteen subjects were secured for each condition. Since it was important to equate the groups for number of trials spent in learning the first list, an attempt was made to secure similar distributions for the four conditions in this respect. That is, an attempt was made to distribute the fast learners and the slow learners equally over the four conditions. This could easily be done since it was possible to assign a subject to a condition after she had completed the original learning. The subjects were all students from elementary psychology courses, and comparable to the subjects in Experiment II.

Results

Trials to Learn Original List.—The mean number of trials required to learn the original list to a criterion of 8 out of 12 right associates for each of the four conditions are as follows: Condition I, 6.86; Condition II, 7.00; Condition III, 7.07; Condition IV, 7.14. The means are very similar, and the slight differences which exist are in

the opposite direction from the expected differences in retention. It is interesting that two more trials were required here, on the average, to learn 8 responses, than in Experiment II, where only five trials were given. The difference may possibly be due to the shorter recall time per syllable in this experiment, or to the influence of social facilitation in the earlier one.

Trials to Learn Interpolated List.—The mean number of trials required to learn the interpolated list to a criterion of 8 out of 12 right responses is presented for each of the four conditions in Table V.

TABLE V

NUMBER OF TRIALS REQUIRED TO LEARN THE INTERPOLATED LIST TO A CRITERION OF 8 CORRECT RESPONSES

Condition	I	II	III	IV
Mean Trials to Learn...	7.43	5.43	5.71	3.93
σ av.....	±1.07	±.65	±.77	±.28

In general, the data corroborate those of Experiment II. There, fewer responses were learned in Condition I than in the other three conditions, which did not differ reliably from one another. Here, more trials are required to learn to the criterion in Condition I than in any other condition. But Condition IV also differs widely from the other three, requiring fewer trials. The difference between Conditions II and III is wholly unreliable.[11] These results tend to confirm the prediction that learning of a second list will be harder as the degree of generalization between the first and the second list increases.

Retention of Original List.—The retention of the original list in the four conditions may be measured in four different ways—by a first recall score, a second recall score, and by two different relearning criteria. In Table VI are presented the results for Recall 1 and 2. The trends in the results of Experiment II are confirmed beyond a doubt, since a gradient in recall from Condition I to Condition IV

TABLE VI

RETENTION OF THE ORIGINAL LIST AS MEASURED BY RECALL 1 AND RECALL 2

Condition	I	II	III	IV
Mean No. Retained, Recall 1.....	2.93	3.00	3.71	6.21
σ av. Recall 1.....	±.45	±.46	±.64	±.34
Mean No. Retained, Recall 2.....	5.00	5.71	7.29	8.36
σ av. Recall 2.....	±.56	±.62	±.67	±.43

[11] The critical ratios ($D/\sigma_{diff.}$) of the differences are: I and II, 1.60; I and III, 1.31; I and IV, 3.18; II and III, .28; II and IV, 2.12; III and IV, 2.18.

is again apparent in both Recall 1 and 2.[12] When the reciprocals of these figures are plotted, the curves are remarkably similar to the ones obtained in Experiment II. In spite of equivalence in degree of interpolated learning, Condition I again shows only slightly worse retention than Condition II. The difference between the two averages is unreliable.[13] The differences between the other conditions are all fairly reliable by one measure or the other. The recall scores thus bear out in every respect the results of the previous experiment and the expectation of the hypothesis that retention should be poorer as degree of generalization between the two lists increases.

Relearning scores in terms of the number of trials required to relearn to one perfect repetition, and the number required to relearn 8 out of the 12 responses may be calculated. Only the latter is correctly called relearning, since the first list was learned originally to this criterion. The results obtained by these two measures are shown in Table VII and Fig. 6. Both relearning measures are consistent with the recall measures in revealing a gradient in retention from Condition I to Condition IV. The difference between Condition I and Condition II is relatively larger as measured by relearning than when it is measured by recall, though it is still not statistically reliable.[14] The slopes of the curves based on relearning are more gradual and nearly linear than those based on recall, but the trend of the results could hardly be more consistently supported.

TABLE VII

RETENTION OF THE ORIGINAL LIST AS MEASURED BY RELEARNING

Condition	I	II	III	IV
Mean Trials to Relearn 8 out of 12	3.79	2.79	2.00	1.36
σ av., 8 out of 12	±.80	±.48	±.41	±.21
Mean Trials to Relearn 12 out of 12	8.29	6.43	5.43	4.79
σ av., 12 out of 12	±1.30	±1.10	±.69	±.53

Analysis of Errors.—Table VIII shows the frequency of occurrence of errors of reversion to the previous list in interpolated learning and relearning. Totals, not averages, are presented, since the ab-

[12] Critical ratios ($D/\sigma_{diff.}$) between alternate conditions are: I and III, .99 first recall, 2.61 second recall; II and IV, 5.61 first recall, 3.54 second recall.

[13] Since interpolated learning was carried to a criterion, and the total interval of interpolation held constant, less time intervened between *cessation* of interpolated learning and *recall* of the original list in Condition I than in the other conditions, giving less opportunity for spontaneous recovery of generalization in this condition. If spontaneous recovery is an important variable, it would then tend to lower the recall scores of the other three conditions as compared with Condition I, thus bringing them all closer to Condition I, where differentiation was most recently achieved.

[14] The critical ratio ($D/\sigma_{diff.}$) is 1.07 for 8 out of 12, and 1.09 for 12 out of 12.

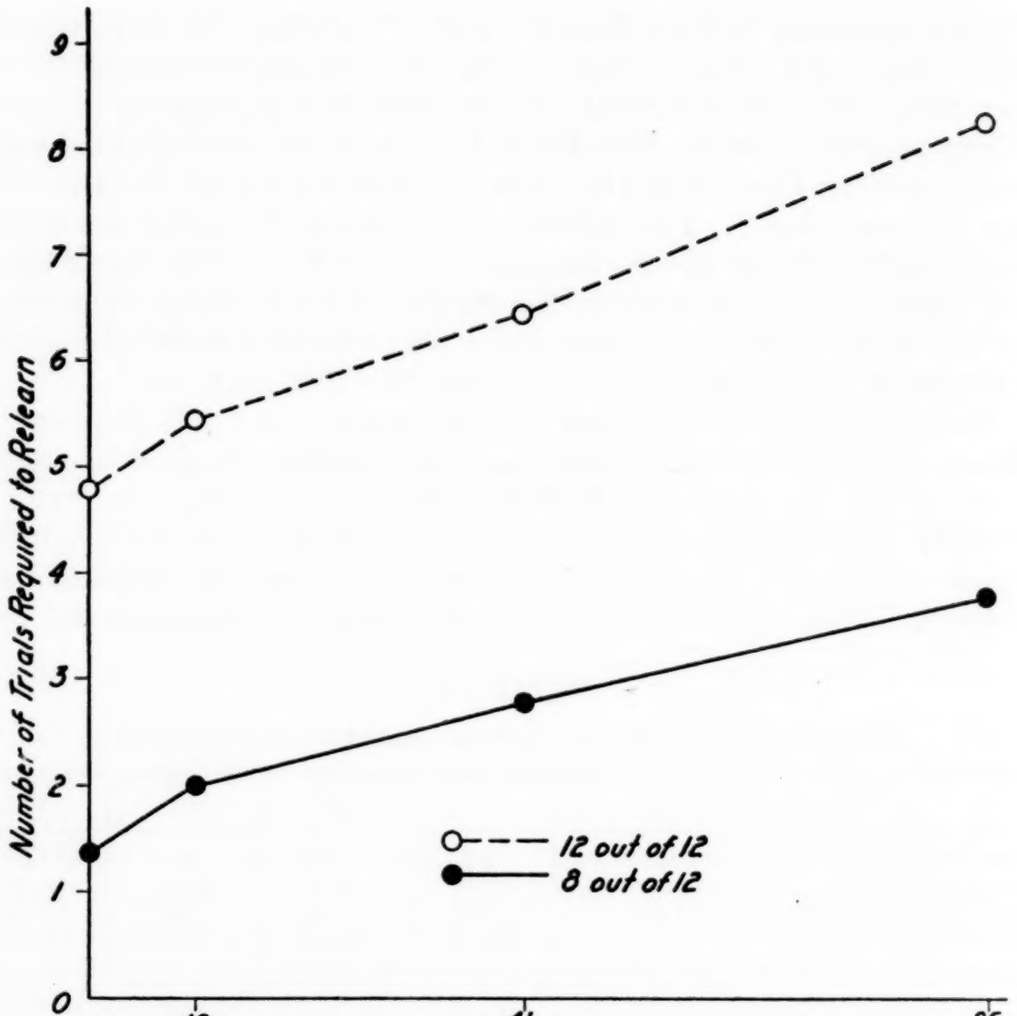


FIG. 6. Trials to relearn List 1 as a function of degree of generalization between the primary and the interpolated lists. The two curves represent number of trials required to relearn to two different criteria.

TABLE VIII

FREQUENCY OF REVERSION ERRORS IN INTERPOLATED LEARNING AND RELEARNING

Condition	I	II	III	IV
Interpolated Learning...	2	4	7	0
Relearning.....	14	23	4	0

solute numbers are very small. The figures are inconsistent, and show no general trend, except that the failure of any such errors to occur in Condition IV is probably significant. The most interesting aspect of these results is the tendency for a larger total number of reversions to the previous list to occur in relearning than in interpolated learning. McKinney and McGeoch (11), in a retroactive

inhibition experiment, also noticed less 'overt' transfer from the original to the interpolated list than from the interpolated list to recall of the original one. The reason for the increase in reversion errors in the latter case is probably not the mere difference in number of repetitions given in the two cases, for in the present experiment they are just about equal, and in McKinney and McGeoch's experiment, only a single recall test was in question. The increase may be related to the fact that a longer time intervened between interpolated learning and recall of List 1 than between original learning and interpolated learning,[15] thereby permitting greater spontaneous recovery of generalization. This explanation requires the assumption that differentiation accomplished *within* a list would have a certain protective function against inter-list generalization. The hypothesis does make such an assumption (4, p. 207).

A fairly high percentage of the errors, as in Experiment II, can be classified as cases of generalization of stimulus members within the list. A total of 51 such errors occurred in each condition on the average during original learning, and 15 during interpolated learning. As in Experiment II, the frequency of intra-list generalization during interpolated learning is consistently less than during original learning in all the conditions, although the original and interpolated lists were learned to the same criterion.

Verbal Reports of Subjects.—Most of the subjects reported that they had not thought of the lists during rest periods. Those who had done so reported that it was impossible really to rehearse, since they had learned to respond in a particular way to visual form cues, and could not think of the responses easily without the cues. In any case, these subjects were distributed among the four conditions, and could not have influenced the trend of the results. Almost all the subjects reported that two responses sometimes occurred to them simultaneously; that there was in such a case a strong tendency to make some response, but that none was made, ordinarily, because the interval was too short for decision. Therefore, some at least of the errors of omission must be due to blocking caused by active competition of responses.

Discussion

The consistency of the results and the complete corroboration of Experiment II when repeated with individual subjects leaves no room for doubt that retroactive inhibition is a function of the degree of

[15] In the present experiment, there was almost always a longer interval between interpolated learning and relearning (especially in Conditions II, III, and IV); and in McKinney and McGeoch's experiment, 9 minutes intervened between interpolated learning and relearning, whereas there was no interval between original and interpolated learning.

generalization between primary and interpolated list. But a consideration of the dynamics of the process is again in order, since this fact does not tell us the actual role of generalization; nor does proof of this role necessarily exhaust the possibilities of explaining decrements in retention.

The process assumed to occur by the present hypothesis is briefly as follows. Generalization may occur from list to list—there will be a tendency for wrong responses to occur in learning a new list or in recalling an old one after interpolation. Active intrusion of wrong responses will thus be a factor leading to decrement. But since these intrusions are infrequent, importance falls on the notion ofcompetition of responses, which may obviously lead to omissions or compromises. Generalized and right responses will tend to block one another. As far as our own data go, active generalization and competitive blocking seem to be sufficient explanations.

A recent article by Melton and Irwin (**12**) furnishes evidence, however, that some process besides the one described above occurs in the typical retroactive inhibition situation. This evidence is in the nature of a curve showing the relationship between the amount of retroactive inhibition and the degree of learning of the interpolated material. When a separate curve was plotted to show the decrement directly attributable to intrusions, the curve increased to a maximum and then decreased as the degree of interpolated learning was increased. This is the relationship predicted for generalization by the present hypothesis (**4**, p. 217). Competitive blocking, according to the hypothesis, should follow the same curve as overt generalization, since strength of the generalized tendency is the crucial factor in both. As differentiation is increased, blocking as well as overt generalization should decrease. But the residual retroactive inhibition, after the curve yielded by intrusions had been subtracted from the total, did not follow the same relationship. Some factor or factors other than generalization and competitive blocking must therefore have contributed to the total inhibitory effect.

The present experiment throws no light on other contributing causes of retroactive inhibition. The writer would like, however, to examine the possibility of extinction or unlearning of the type found in the conditioned response situation as a factor, since such an hypothesis has been attributed to her (**12**). The conditions for extinction are present during interpolated learning, in so far as differential reinforcement takes place. If the subject gives a generalized response, it will not be reinforced and another response will be indicated as the correct one. In this situation, therefore, the generalized response may be extinguished through differential inhibition. But this response itself is not generally inhibited, of course. When

a salivary response to a bell is extinguished, there is not inhibition of salivation in general, but only of salivation in that situation or similar ones. If a parallel is drawn in the present situation, the generalized response has been extinguished as far as the stimulus in the interpolated list is concerned; another response has been learned as the positive one for that stimulus. But has the extinguished response been inhibited for its own stimulus of the first list—its positive stimulus? Generalization of inhibition is a well-known phenomenon (6, p. 178 ff.), but to the writer's knowledge it has not been demonstrated in a discrimination situation. There is not even an analogue, then, to support the required assumption that the inhibition of generalized responses during interpolated learning will itself generalize to the positive stimulus items of the first list. Suppose such an assumption is made, however. It is still a fact that very few generalized responses actually occurred overtly during interpolated learning, so that the opportunities for inhibition of generalized responses were few. It might be argued that non-reinforcement of *any* wrong response will produce inhibitory effects, or that *implicit* generalized responses might be extinguished. As for the effects of extinguishing *any* wrong response, it seems clear that such extinction could not contribute to retroactive inhibition, because unless the punished errors were due to generalization, they would not be at all relevant to recall of List 1. It would be necessary to assume a concept of 'free-floating' inhibition to find a way for such non-reinforcement to affect recall of List 1. As for the possibility that inhibition may result from the non-reinforcement of implicit generalized responses, the following arguments present themselves. If a subject thinks of a generalized response, but does not speak it, one of two things is probably happening. If it is the only response he thinks of, he does not speak it because he already knows it to be wrong, in which case he *has differentiated* it. In the writer's opinion, the essential condition for retroactive inhibition is thereby removed. If it occurs along with another possible response, the situation which has been referred to as competitive blocking results. In so far as the competing generalized response is non-reinforced, it may be extinguished in that situation. The possibilities of extinction are therefore limited to overt generalized responses and competing generalized tendencies. Since in both these cases it is necessary to assume generalization of differential inhibition to the *positive* stimulus, the theoretical structure required is elaborate and not too plausible.

Other theoretical possibilities for retroactive inhibition must be kept in mind at the present time, especially because it has been shown to occur in certain situations which are difficult to analyze in the present terminology. Retention of forms in primary memory, espe-

cially, may be mentioned. Here Gibson and Raffel (5) have shown retroactive inhibitory effects for the reproducing of a given visual form when it has been succeeded by other forms. The retroactive inhibitory effect increases progressively with an increase in the number of different interpolated forms. Subjects reported a 'blotting-out' of the memory image as new forms made their appearance. In this situation it seems likely, therefore, that the principal locus of the decrement occurred while the interpolated impressions were taking place, and that interference at the time of recall played a small part. Nor could inhibition resulting from differential reinforcement be involved. It is possible that when 'content' retention is primarily involved, as in this situation (and presumably to some degree in any learning situation), a weakening of the original impression during interpolation is an important factor. The weakening effect does not necessarily depend upon the interruption of perseveration, as has often been assumed; it might take the form of distortion or deletion of the original 'trace.'

In conclusion, the writer feels that a determination of the role and weight of the various factors suggested can only be made after careful consideration has led to prediction of the effects which they should produce when the usual dimensions (similarity, degree of learning, etc.) are varied in the retroactive inhibition situation.

Conclusions

1. When syllable responses have been learned to a list of stimulus forms, variations from these forms may be shown to generalize with the original forms, since the variations are responded to in a test series as if they were the originals. Gradients of generalization in terms of frequency of response to the variations may be demonstrated, and such gradients correspond with gradients of perceptual similarity.

2. As the degree of generalization between corresponding stimulus members of a first and a second list is increased, when responses in the two lists are different, there is a tendency for the second list to be harder to learn. The second list is hardest to learn when the stimulus members of the two lists are identical.

3. The degree of retroactive inhibition of a first list varies directly with the strength of the tendency for stimulus members of an interpolated list to generalize with those of an original list.

4. Errors of overt generalization, or reversion to the previous list occur during both interpolated learning and relearning of the primary list. There is some tendency for frequency of such errors to vary with the degree of generalization between the two lists.

5. Errors of intra-list generalization occur during the learning of

both the primary and the interpolated list, but such errors decrease considerably from primary to interpolated learning.

(Manuscript received September 6, 1940)

References

1. Britt, S. H., Retroactive inhibition: a review of the literature, *Psychol. Bull.*, 1935, **32**, 381–440.
2. Gibson, E. J., *A systematic application of the concepts of generalization and differentiation to verbal learning*, Dissertation on file in Yale University Library.
3. Gibson, E. J., Retroactive inhibition as a function of degree of generalization, *Psychol. Bull.*, 1938, **35**, 626.
4. Gibson, E. J., A systematic application of the concepts of generalization and differentiation to verbal learning, *Psychol. Rev.*, 1940, **47**, 196–229.
5. Gibson, J. J., and Raffel, G., A technique for investigating retroactive and other inhibitory effects in immediate memory, *J. gen. Psychol.*, 1936, **15**, 107–116.
6. Hilgard, E. R., and Marquis, D. G., *Conditioning and Learning*, Appleton-Century Co., New York, 1940, pp. 429.
7. Hull, C. L., The meaningfulness of 320 selected nonsense syllables, *Amer. J. Psychol.*, 1933, **45**, 730–734.
8. McGeoch, J. A., The influence of degree of learning upon retroactive inhibition, *Amer. J. Psychol.*, 1929, **41**, 252–262.
9. McGeoch, J. A., The influence of degree of interpolated learning upon retroactive inhibition, *Amer. J. Psychol.*, 1932, **44**, 695–708.
10. McGeoch, J. A., and McGeoch, G. O., Studies in retroactive inhibition: X. The influence of similarity of meaning between lists of paired associates, *J. exper. Psychol.*, 1937, **21**, 320–329.
11. McKinney, F., and McGeoch, J. A., The character and extent of transfer in retroactive inhibition: disparate serial lists, *Amer. J. Psychol.*, 1935, **47**, 409–423.
12. Melton, A. W., and Irwin, J. M., The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses, *Amer. J. Psychol.*, 1940, **53**, 173–203.
13. Thorndike, E. L., A note on assimilation and interference, *Amer. J. Psychol.*, 1937, **49**, 671–676.
14. von Restorff, H., Analyse von Vorgängen im Spurenfeld. I. Über die Wirkung von Bereichsbildungen im Spurenfeld, *Psychol. Forsch.*, 1933, **18**, 299–342.
15. Yum, K. S., An experimental test of the law of assimilation, *J. exper. Psychol.*, 1931, **14**, 68–82.

PREPARATORY SET (EXPECTANCY)—FURTHER EVIDENCE OF ITS 'CENTRAL' LOCUS*

BY O. H. MOWRER

Harvard University

The Problem

In a recent paper, Mowrer, Rayman, and Bliss (11) have reported experimental evidence which they interpret as indicating that "preparatory set (expectancy, attention) is mediated principally by neural rather than by neuro-muscular mechanisms" (p. 371). These writers concede that muscular adjustments do, in fact, often accompany sets, but they maintain that the latter can exist and vary independently of the former. This position is in opposition to the pristine behavioristic view that preparatory set must, like all other psychological phenomena, necessarily have motor correlates. The principal evidence which Mowrer, Rayman, and Bliss give in support of their hypothesis is the finding that if human subjects are instructed to make a simple finger-withdrawal response as quickly as possible whenever *either* a light or a tone is presented, reaction time is much faster to the one type of stimulus than to the other, depending upon which is the more confidently 'expected' at the moment. More specifically, it was found that if several lights are presented at regular (six-second) intervals and if a change to a series of tones then occurs without warning, the reaction time to the first of the series of tones is abnormally long. Similarly, if a series of lights follows unexpectedly upon a series of tones, the latency of the response to the first light is likewise much increased. Since the same identical response, namely, finger-withdrawal, was made to both types of stimuli, it seemed highly improbable that differential muscular adjustments on the part of the responding (right) hand and arm could account for the observed results. In other words, optimal tonic and postural adjustment of the hand and arm for responding to one type of stimulus seemed to provide an equally good preparation for making the same reaction to the other type of stimulus; and since gross variations in reaction time were observed (indicating differential preparedness), it was concluded that these variations could not be correlated with muscular changes in the responding member but must be due to factors operating elsewhere within the organism. Since it also seemed unlikely that differential motor adjustments of the optic and auditory sense organs could account for the results obtained, it was concluded

* This investigation was carried out at the Institute of Human Relations, Yale University.

that these results were most probably due to shifts in set which were mediated by purely neural, or 'central,' factors.

This interpretation has been criticized by Freeman (4), who agrees that the results of the study cited satisfactorily demonstrate "that the responsible factors are [not] the covert set of the muscles to be involved in the overt reaction—since finger preparation *per se* is the same for both stimuli" (p. 624) and that "there need be no differential pattern of antecedent tension in the reacting finger for different expectancies" (p. 625). However, Freeman insists upon the thoroughgoing behavioristic view that "both central 'neural' and peripheral 'muscular' conditions are involved in any set" (p. 623) and hypothecates that since the 'muscular' factor is not demonstrable at the motor end of the response arc in the experiment cited, it must be localizable in the adjustor mechanisms of the receptors. Freeman says:

> Mowrer, Rayman, and Bliss admit that their experiments would have permitted differential motor adjustments of the receptive sense organs, but they airily dismiss the possibility of their occurrence as "very remote . . . since movement of the ears could not be made similarly [as might variations in visual sense organ adjustment] to account for the increased latency of the reactions to the first of a series of tones following a series of lights." It is precisely because of this unchecked assumption that their 'demonstration' of the set-locus breaks down. The lengthened reaction time to first presentations of tone stimuli may have been due to either the unprepared state of the *tensor tympani* or to the absence of general motor adjustments favoring hearing. . . . In my opinion, Mowrer, Rayman, and Bliss's ruling out of motor factors as component determinants of set, by the method of proclamation rather than by adequate control, leaves us with no alternative explanation other than the soul hypothesis masquerading in intra-neural false whiskers (pp. 623–624).

The issue is clear-cut. Freeman, accepting Bok's concept of the 'reflex circle' ('backlash proprioception') as the necessary basis of preparatory set, holds that *both* central and muscular factors must be involved; if the muscular factor is not discoverable at the motor end of the response arc, then it must exist at the sensory end. (The alternative hypothesis, that set may be mediated by purely central mechanisms, is based upon the concept of 'reverberating,' or 'after-discharge,' neural circuits of the type described by Lorente de Nó (7)[1] and holds that preparatory set may, on the contrary, involve only central factors.[2]

[1] In the study by Mowrer, Rayman, and Bliss, these writers did not make explicit reference to reverberating neural circuits as the central basis of preparatory set but such mechanisms were tacitly assumed by them. Freeman dismisses this possibility of accounting for the results in question by saying that "neural traces of the different stimuli . . . are relatively short-lived" (p. 624). What is the evidence for such a statement? Precisely what is meant by 'relatively short-lived?' Excitation from the vestibular receptors can reverberate in the central nervous system for 30 or more seconds (8, 9). Dusser de Barenne and McCulloch (2, 3) have reported that artificial stimulation of the monkey cortex is capable of producing electrical after-discharge phenomena lasting 'one to several minutes,' even 'after deep undercutting of cortex,' *i.e.*, after severance of the neural pathways communicating with the skeletal musculature. Pavlov (12) found that dogs can establish trace conditioned salivary responses after delays of as long as half an hour. What are the experimental findings which prove that excitation originating in the visual

Freeman has proposed one method whereby these two hypotheses can be differentially tested. He says: "A crucial experiment for demonstrating the central locus of set would be one which showed (by action potentials, eye movement records, or otherwise) that the incipient motor adjustments forming the non-focal parts of the preparatory tonic background did not shift appropriately during the building up of different stimulus expectancies" (p. 626). This would be a feasible approach to the problem were it not for the technical difficulties involved in detecting incipient activity in such small and inacessible muscles as those associated with sense organs such as the eye and ear and for the fact that, in the event of negative results, advocates of the Freeman position could always maintain that activity was indeed present in the muscles but that the methods employed were not sufficiently refined to demonstrate it. The experiments which will here be reported follow a different course and were designed to avoid these pitfalls.

Procedure and Results—Section A

One possibility of meeting the objection that the findings reported by Mowrer, Rayman, and Bliss were due to differential sense-organ adjustment rather than to central factors would be to repeat their experiment but to restrict stimulation to a single sensory modality and to employ therein two stimuli which are perceptibly different but which do not call for differential sense-organ adjustment. If, under these conditions, each unexpected change from one type of stimulus to the other brought an increase in reaction time, this finding would argue strongly against the view that differential sense-organ adjustments are responsible for this effect. On the other hand, negative results, *i.e.*, failure to obtain a reaction-time increase under these conditions, would be ambiguous, neither proving nor disproving either of the alternative hypotheses discussed above. Despite the fact that an experiment of this kind would be significant only in the event of positive results, it nevertheless seemed worth performing. It was accordingly carried out in three forms, as follows.

Experiment I

Each of the 20 college men employed in this experiment was installed in a comfortable chair in a sound-proofed dark room[3] and was shown a round milk-glass window, four cm in diameter,

and auditory receptors cannot persist in the central nervous system of human beings for six seconds?

[2] Again it should be stressed that this hypothesis does not hold that muscular factors are never involved in preparatory sets but simply that such factors are not a prerequisite for the existence of sets.

[3] The same general experimental set-up employed by Mowrer, Rayman, and Bliss was used in the present investigation. The only major changes were in the types of stimuli presented.

which was centrally located on the near side of a small box situated approximately 96 cm directly in front of the subject at the level of his eyes. Each subject was then told that from time to time the letter 'S' or the letter 'T' (both of which were 18 mm high and could be projected on exactly the same area of the milk-glass) would appear in this window and that it was the subject's task to make these letters disappear as soon as he saw either of them. This could be done by releasing a telegraph key which was then shown to the subject and which he was instructed to hold down continuously, except when he saw one of the letters. Finally, subjects were admonished to make their responses to the letters as quickly as they possibly could and to keep their eyes open and directed toward the milk-glass window throughout the experiment,[4] which, they were told, would take a little less than four minutes.

The order of presentation of the two visual stimuli used in this experiment was as follows: 7 T's, 11 S's, 3 T's, 9 S's, 4 T's, and 1 S. The average reaction times of all 20 subjects to these 35 successive stimulus presentations are represented graphically in Fig. 1. It will

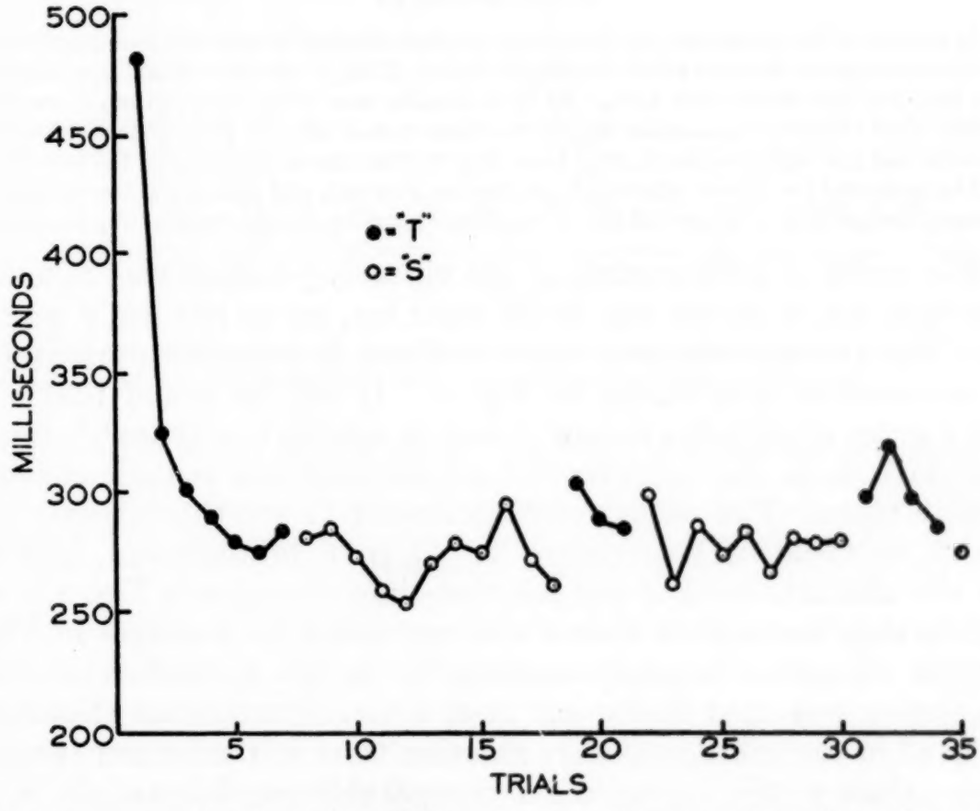


FIG. 1. Average reaction times of 20 college men to irregular series of visual stimuli, consisting of the letter 'T' and the letter 'S.'

be seen that at no point did the shift from a series of one type of stimulus to a series of the other type produce a marked increase in reaction time. Although in four out of five of these shifts an increase

[4] In preliminary experimentation a luminous spot was provided just above the exposure window as a means of assisting fixation during the intervals of total darkness between presentations of the letters. However, because of adaptation and after-image effects, this arrangement proved confusing and was discontinued. The intensity of the lights used to project the letters was relatively weak so that after-images of the latter were not conspicuous. The filaments of the two flashlight bulbs which provided the illumination required for this purpose were kept 'soaked' so that their 'on' latency was relatively brief.

in reaction time did occur, these increases are statistically unreliable. The largest of these increases was 41 milliseconds, but the critical ratio of the difference represented by this figure is only 0.49, which means that there are about 30 chances in 100 that it was accidental. There are some indications that if a great many more subjects had been run in this experiment reliable increases in reaction time might have been obtained. But the comparatively small number of subjects here used suffices to show that such differences, if real, are very small in comparison with those found by Mowrer, Rayman, and Bliss in their investigation. Such slight differences did not seem to warrant the use of more subjects.

Experiment II

At the end of the procedure just described, the experimenter entered the sound-proofed room and told each subject that two small electrically-driven (Hull, 1) vibrators would now be attached to his legs, one just above each knee. As these devices were being strapped on, it was further explained that vibratory stimulation on the two knees would take the place of the letters 'S' and 'T' which had just been employed, that from time to time one or the other of the two vibrators would be activated (by 4 volts of 60-cycle alternating current), and that it was the subject's task whenever this occurred to release his key as quickly as possible, thereby terminating the vibration.

The order of presentation of the vibratory stimuli was as follows: 2 on right leg, 6 on left leg, 10 on right leg, 10 on left leg, 7 on right leg. The average reaction times to these 35 stimulus presentations are represented graphically in Fig. 2. It will be noted that shifts from a series of stimulus presentations on one leg to a series of stimulus presentations on the other leg did not produce any reliable change in reaction time. The results of Experiment II are thus strictly comparable to those of Experiment I. It is to be observed, however, that the absolute level of the reaction-time averages in Fig. 2 is very considerably lower than that of the reaction-time averages in Fig. 1; but this difference is easily explainable on the grounds that visual stimulation was used in the one case, whereas cutaneous stimulation was used in the other, and that reaction time to the former is always greater than to the latter under comparable conditions. It is also noteworthy that the "practice effect" is much more precipitate in Fig. 2 than in Fig. 1. This difference is presumably due to the fact that in Experiment I the subjects had to learn the rate at which the successive stimuli occurred, *i.e.*, they had to learn to "time" their set. Since the same rate of stimulus presentation had been used in Experiment I as was used in Experiment II, learning of this kind was less necessary in the latter case.

Experiment III

In the next experiment, which was conducted with the same subjects, two visual stimuli were again employed, as in Experiment I, but instead of being differentiated on the basis of form, they consisted of a red and a green light which could independently be made to suffuse the whole of

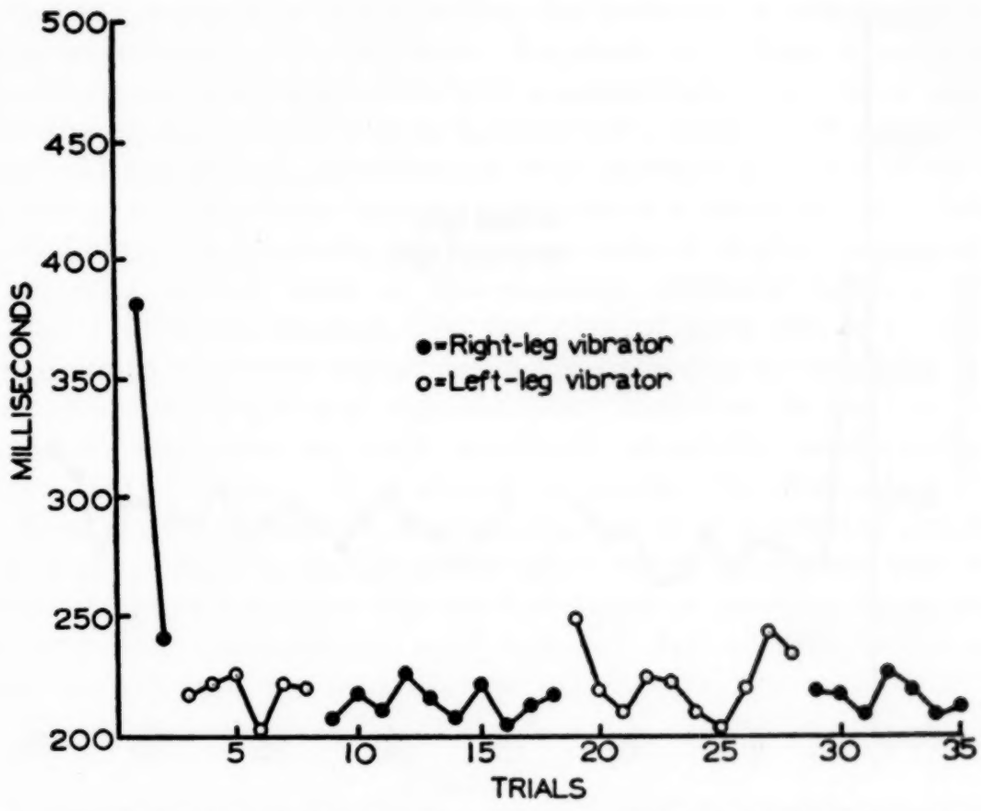


FIG. 2. Average reaction times of 20 college men to irregular series of tactual stimuli supplied by two mechanical vibrators, one located just above either knee.

the round milk-glass window described above.[5] The order in which these two visual stimuli were presented was as follows: 12 red lights, 2 green lights, 8 red lights, 7 green lights, 5 red lights, and 1 green light. The average reaction times to these 35 stimulus presentations, which are graphically represented in Fig. 3, show no significant change with shifts from one type of stimulus to the other. These results are therefore parallel to those obtained in Experiments I and II.

The absolute level of the reaction time averages in Fig. 3 is about the same as that of the reaction time averages in Fig. 1. It is interesting to note that in both cases there was a gradual increase in reaction time after the first few trials, due, presumably, to the operation of fatigue and possible loss of motivation. Why a similar effect was not obtained in Experiment II, employing tactual stimuli, is not apparent. In Fig. 3 the same precipitousness of the initial practice effect is to be noted as in Fig. 2.

The results obtained in the three experiments just reported must be regarded as wholly ambiguous insofar as the main purpose of this study is concerned. That is to say, they neither support nor oppose the hypothesis of Mowrer, Rayman, and Bliss that preparatory set may be mediated by purely neural factors nor do they either support

[5] The red light was produced by means of a Rosco gelatine filter, No. 110 (Fire Red), and the green light by means of a Rosco gelatine filter, No. 128 (Medium Green). These two colors were roughly equated for brightness and were both sufficiently subdued so that after-images from a brief exposure were hardly perceptible.

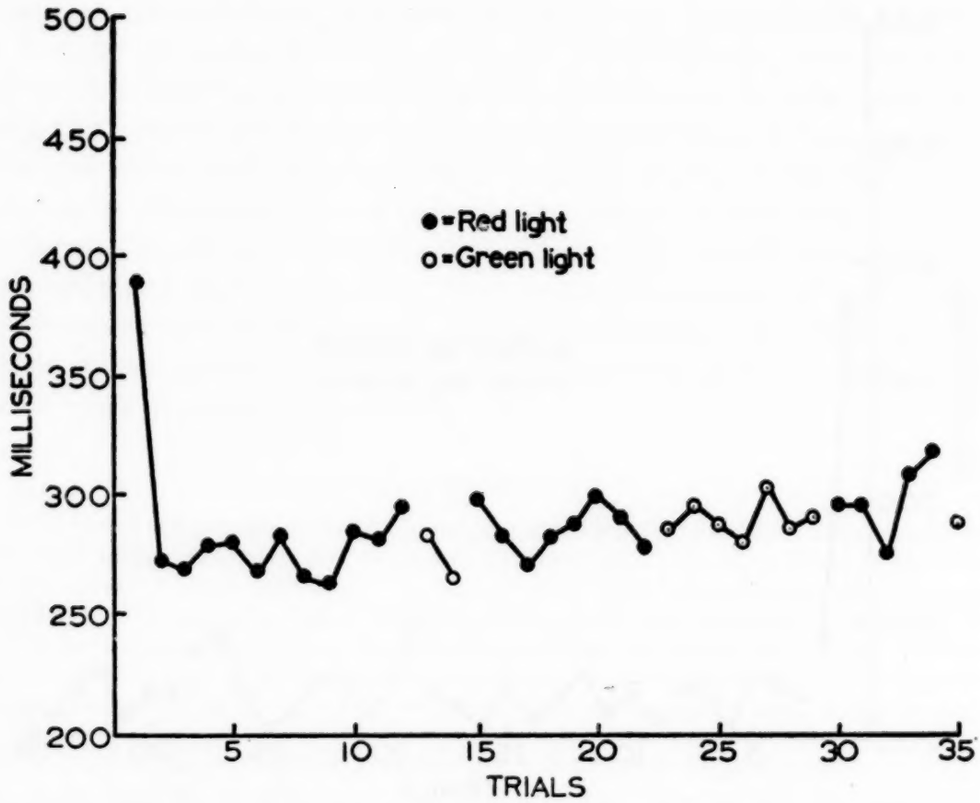


FIG. 3. Average reaction times of 20 college men to irregular series of visual stimuli, consisting of a red light and a green light.

or oppose the hypothesis of Freeman that set necessarily involves continual 'backlash' stimulation from muscular mechanisms. In the event of positive findings in Experiments I, II, and III, the first of these hypotheses would have been decisively confirmed. Lacking definitive results from these experiments, it was accordingly necessary to find other approaches to the problem.

PROCEDURE AND RESULTS—SECTION B

The fact that unexpected shifts from one type of stimulus to another, within the same sensory modality, did not (in Experiments I, II, and III) result in heightened reaction latencies of the type reported by Mowrer, Rayman, and Bliss in their work with visual and auditory stimuli suggested that this effect may be dependent upon the crossing from one sensory modality to another. It therefore seemed desirable to check the Mowrer-Rayman-Bliss findings, using, however, auditory and cutaneous stimuli instead of auditory and visual stimuli. This arrangement provided cross-modality stimulation but at the same time minimized the possibility that differential sense-organ adjustment could account for the results in the event that they were positive. It was, as Freeman rightly observed, not

beyond the realm of possibility that the increases in reaction latency, which were observed by Mowrer, Rayman, and Bliss when a series of auditory stimuli were followed unexpectedly by visual stimuli, were due to the subject's eyes having been caught 'off guard,' as it were. Although less plausible, it was likewise not impossible that the increases in reaction latency noted when a series of visual stimuli was followed unexpectedly by auditory stimuli might have been due to the ears having been at the moment similarly unprepared for optimally efficient hearing. (In the case of both the eyes and the ears, definite muscular adjustments are known to be essential for the most effective reception of stimulation. However, in the case of the cutaneous receptors no such auxiliary muscular mechanisms are known to the writer. If it should be found, therefore, that an unexpected shift from a series of auditory stimuli to a series of cutaneous stimuli resulted in a significant increase in reaction time, this would constitute good evidence that such changes in reaction time (and in the attendant preparatory sets) are not due to differential sense-organ adjustment of a muscular nature.)

Experiment IV

In order to submit this deduction to an empirical check, the simple reaction times of 30 college men (not used in the preceding experiments) were tested in the manner just described, with auditory and vibratory stimuli. The vibratory stimulus, which was applied above the knee of the right leg, was of exactly the same type as employed in Experiment II, *i.e.*, it was derived from a Hull vibrator, driven by 4 volts of 60 cycle current. The auditory stimulus consisted of a relatively pure tone, 800 d.v./sec in pitch, which was delivered to the subject through headphones at an intensity of about 70 decibels above threshold. Although the vibrator was practically noiseless, a 48 decibel, 60 cycle masking tone was nevertheless sounded continuously in the headphones during experimentation in order to reduce the possibility of the subject's responding to the auditory, as well as to the tactual, stimulation given off by the vibrator. The order of presentation of the two types of stimuli was as follows: 7 tones, 11 vibrations, 3 tones, 9 vibrations, 4 tones, and 1 vibration.

The average reaction times of all subjects to each of these 35 stimuli are presented in Fig. 4. The average reaction time to the 8th stimulus, which was auditory, was 243.2 ms, whereas the average reaction time to the 9th stimulus, which was vibratory, was 421.2 ms. The difference between these two figures is 178.0 ms, with a critical ratio of 5.52, which establishes the reliability of this difference with virtual certainty. The average reaction time of all subjects to the 18th stimulus, which was vibratory, was 255.3 ms, and the average reaction time to the 19th stimulus, which was auditory, was 311.1 ms. The difference is 55.8 ms., with a critical ratio of 2.44, which means that the chances are about 99.3 in 100 that this difference did not occur through accident but was due instead to the change from one type of stimulus to the other. Between the 21st and 22nd, the 30th and 31st, and the 34th and 35th stimuli there were similar

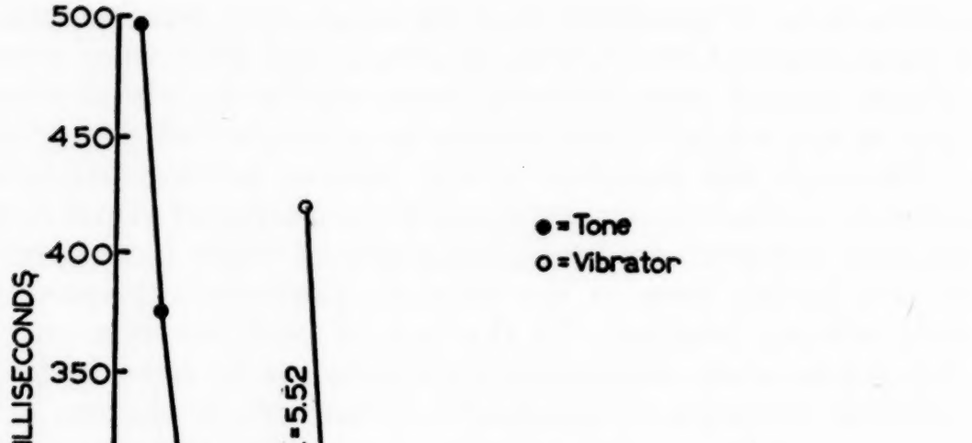


FIG. 4. Average reaction times of 30 college men to irregular series of auditory and tactual stimuli (tone and mechanical vibrator).

changes from vibratory to auditory stimuli or *vice versa*. As Fig. 4 shows, all of these changes were accompanied by increases in average reaction time, but the magnitude of these increases is small and not highly reliable statistically.

The foregoing results indicate that an unexpected change from a series of auditory stimuli to vibratory stimuli is capable of producing a reliable increase in reaction time, and since there seems to be no way of accounting for this on the basis of inadequate muscular preparedness of the tactile receptors, these results are taken as confirmatory of the hypothesis that changes can occur in preparatory set (as measured by quickness of reaction time) without concomitant muscular changes, either at the sensory or at the motor end of the response arc.

Two additional observations may be made in connection with the data represented in Fig. 4, the first of which is the fact that the magnitude of the reaction-time increases which occurred with each change from one type of stimulus to the other tended to become progressively smaller. Whether the magnitude of these increases would have eventually become zero if additional series of vibratory and auditory stimuli had been presented cannot be said on the basis of the present

data, but the indications are that such might have been the case. This observation stands in contrast to the results reported by Mowrer, Rayman, and Bliss. Using visual and auditory stimuli, these writers found no consistent tendency for the magnitude of reaction-time increase to diminish with successive changes from one type of stimulus to the other. It is true that their data cover only four such changes and that an extension of their observations to additional changes might have revealed a 'learning effect' of the kind which seems to be noticeable in Fig. 4. However, it is also true that the writers cited used considerably longer series of homogeneous stimuli (27 tones, 9 lights, 13 tones, 21 lights, [and] 1 tone) than were employed in the present study. This may account for the larger and more reliable reaction-time increases noted by these writers.

The practice adopted in the present study of having some very short series of homogeneous stimuli (as few as 2 or 3 stimuli in some cases) was based on the assumption that after several relatively long alternative series of homogeneous stimuli, a change from one to the other type of stimulus at the end of a very short series would be quite as effective in lengthening reaction time (*i.e.*, would be quite as 'unexpected' by the subject) as would changes after the much longer series. Although this assumption would seem to be in keeping with a common sense, intuitive view of the matter, there is now evidence for suspecting that it is not true. Mowrer, Rayman, and Bliss (11) have reported and Winer (13) has confirmed that if two types of stimuli are presented in simple alternation several times and if one type of stimulus is then unexpectedly presented twice in succession, there is ordinarily no increase in reaction time. Although no satisfactory explanation can at present be given for this fact, it seems evident that reaction-time increases of the type under discussion here can be reliably obtained only when relatively long series of one type of stimulus presentations alternate with relatively long series of the other type of stimulus presentations. It is regrettable that this discovery was not taken into account in designing the experiment reported in this section of the present study. The results obtained would have conjecturally been of a more dramativ nature, although they suffice as they are to provide strong support for the Mowrer-Rayman-Bliss hypothesis that preparatory set is not inevitably dependent upon muscular factors.

The other incidental finding to which attention should be called in connection with the data presented in Fig. 4 is that the absolute level of the average reaction times to the two series of vibratory stimuli is considerably higher than that of the vibratory reaction-time averages shown in Fig. 2. Although the statistical reliability of this difference has not been determined, its magnitude and relative

uniformity from trial to trial suggest that it is probably a valid phenomenon. If this be the case, the question arises: Why does the presentation of series of vibratory stimuli to the right leg result in longer reaction latencies when they are separated by series of auditory stimuli (as in Fig. 4) than when separated by series of comparable vibratory stimuli presented to the left leg (as in Fig. 2)? The answer to this query seems to be provided by an experimental finding reported by Mowrer, Rayman, and Bliss (11). These writers found—and Winer (13) has confirmed—that if subjects are instructed to be prepared to react to either tones *or* lights, their reaction times to a long, uninterrupted series of tones will, at least in the beginning, be materially greater than if the subjects are presented with an identical series of tones but with no intimation that they should also be prepared to react to lights. The writers cited refer to the former of these conditions as involving 'divided attention' and to the latter as involving 'unified attention.' If it can be assumed that in Experiment II in the present investigation the subjects regarded vibration on the right leg as essentially the *same* stimulus as vibration on the left leg, then they can be said to have been in a state of 'unified attention,' with a consequent reduction of absolute reaction time. On the other hand, the subjects in Experiment IV would certainly not have regarded vibration on the right leg as the same stimulus as the 800 v.s./sec tone which was presented to them through the headphones which they wore. It is possible that with practice the life-long habits of discrimination between auditory and tactual stimuli might be sufficiently broken down to permit a kind of psychological equating of these two types of stimuli, for, under the conditions of the experiment, they were indeed equivalent in that both produced the same response. As a matter of fact, it may have been just such a process as this which was responsible for the progressive diminution of the reaction-time increase with successive changes from one type of stimulus to the other which is noted in Fig. 4. This hypothesis that change from a series of one type of stimulus to a series of another type of stimulus will or will not result in a heightened reaction time, depending upon whether the two types of stimuli are equated or differentiated by the subject, can, fortunately, be tested by repeating the type of experiment reported by Mowrer, Rayman, and Bliss and carrying it through a great many more changes than did these writers. If the hypothesis is correct, there should ultimately be a disappearance of the reaction-time increase with change from one type of stimulus to another.

If the failure in Experiment II to obtain an increase in reaction time with changes from vibratory stimulation of the right leg to vibratory stimulation of the left leg, and *vice versa*, was, as here con-

jectured, due to the fact that the subjects were really reacting to these two spatially distinct stimuli as to one and the same stimulus, a similar explanation can be offered for the negative results obtained in Experiments I and III. In the first of these experiments, it is entirely possible that the subjects were reacting, not to the particularities of form of the stimulus 'S' and of the stimulus 'T' but to the feature common to both, namely, illumination. In Experiment III it is likewise possible that the subjects were not reacting on the basis of the hues 'red' and 'green' but instead, to the feature common to both, again illumination.

These speculations raise a number of questions concerning the relation between perceptual discrimination and identification, on the one hand, and response diversity and identity, on the other hand. These cannot, however, be relevantly treated in the present context but provide some suggestive leads for further experimentation.

Experiment V

It will be recalled from an earlier section of this paper that Freeman, in criticizing the Mowrer-Rayman-Bliss hypothesis concerning the 'central' locus of preparatory set, has taken the position that the experimental findings reported by these writers (and confirmed by Winer and the experiment just described) is due to "differential motor adjustments of the receptive sense organs." In the first three experiments reported in this study an attempt has been made to eliminate the possibility of differential sense-organ adjustments by using in each experiment two stimuli within the *same* sensory modality. Negative results were obtained in the sense that under these conditions there was no increase in the subjects' reaction times with changes from one type of stimulus to the other; but, as already pointed out, this does not necessarily prove that this effect was absent because differential sense-organ adjustment had been eliminated. It could equally well have been due to the fact that cross-modality stimulation had also been excluded. In Experiment IV this latter supposition tends to be confirmed by the finding that although there are no known muscles which specifically adjust the tactual sense organs, reaction-time increases can nevertheless be produced by changes from auditory to tactual stimulation.

If, as Freeman proposes, the increase in reaction time reported by Mowrer, Rayman, and Bliss is due to differential sense-organ adjustment, then it should be possible to augment this effect by deliberately increasing the necessity of differential sense-organ adjustment. In Experiment I and Experiment III it was found that reaction-time increase did not occur with shifts from a series of one type of stimulus

to another, providing that both stimuli were visual and did not call for differential receptor adjustments. If the negative results obtained in these experiments were due to the absence of differential receptor adjustments, rather than to the absence of cross-modality stimulation, then it should be possible to obtain positive results, *i.e.*, sizeable reaction-time increases with change from one type of stimulus to the other, providing that these stimuli, even though within the same modality, be such as to require differential sense-organ adjustment for optimal reception.

In order to test this deduction, the same 30 subjects which served in Experiment IV were submitted to the following procedure. Immediately after completing the Experiment-IV procedure, each subject was given a few minutes of rest. Each subject was then shown two 3-watt neon lamps located at a distance of six feet and at the level of the eyes on the opposite wall of the experimental room. These lights were separated from each other by a distance of 60° of visual angle, which meant that in order for the subject to look from one to the other he had to execute a major eye movement. The subject was told that sometimes one of these lamps would flash on and sometimes the other and that it was his task to turn off whichever one came on as quickly as possible by releasing the telegraph key located on the right arm of his chair, just as he had terminated the tone and the vibrator in the preceding experiment.

As in the preceding experiments, 35 stimulations were presented in all, but in the following order: 2 lights at right, 6 lights at left, 10 lights at right, 10 lights at left, 4 lights at right, and 3 lights at left. As will be seen by reference to Fig. 5, at no point did the shift from

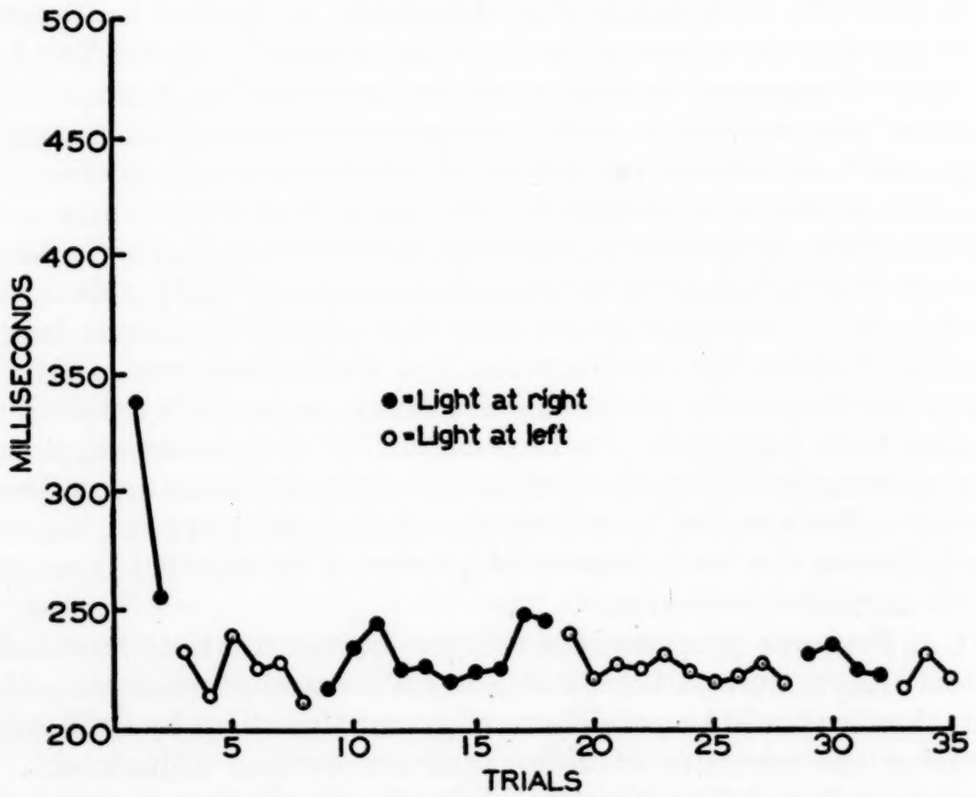


FIG. 5. Average reaction times of 30 college men to irregular series of visual stimuli, consisting of two neon lamps separated by 60° of visual angle.

one type of stimulus to the other elicit a significant increase in reaction time. This finding is interpreted as indicating that differential ocular adjustment, at least of the kind demanded in the present experiment, is of little or no relevancy in explaining the reaction-time increases reported by Mowrer, Rayman, and Bliss as occurring with unexpected shifts from auditory to visual stimulation and *vice versa*.

This finding is actually somewhat surprising in view of the early work of Hall and von Kries (**6**), which suggested that the latency of a simple reaction to visual stimulation tends to be longer if a peripheral rather than a central region of the retina is involved. If, as a result of a series of presentations of the light located at the subject's right in the present experiment, his eyes tended to remain fixated in the general direction of this light, this would mean that when the light at the left was presented, its image would first fall on a decidedly peripheral portion of the retina. Why stimulus changes of this kind did not, as the work of Hall and von Kries would have led one to expect, cause the subjects to show a heightened reaction time when their eyes were thus caught 'off guard' is not at present clear. But whatever the explanation of this contradiction,[6] the present results constitute additional evidence for the view that the findings reported by Mowrer, Rayman, and Bliss were dependent, not upon differential sense-organ adjustments, but upon the development of differential neural sets, which, as now seems likely, are associated with some peculiarity of cross-modality stimulation.

Comparison of the results obtained in Experiments I and III (see Fig. 1 and Fig. 3) with those obtained in Experiment V (Fig. 5) reveals that the general level of reaction time was approximately 100 ms. lower in the latter case. Since visual stimulation was involved in all three experiments, this difference might be assumed to have undue significance. It will be recalled, however, that in Experiments I and III the visual stimulation was provided by small tungsten lamps with 'soaked' filaments, whereas in Experiment V stimulation was supplied by neon lamps. Although the 'soaking' considerably reduced the initial, or 'on,' latencies of the tungsten lamps, they were probably still considerably slower in flashing on than were the neon lamps. This difference in lamp-latency presumably accounts for the difference in general reaction-time average just noted.

Discussion

In his criticism of the experiments reported by Mowrer, Rayman, and Bliss (**11**), Freeman (**4**) has not only questioned the validity of

[6] It is conceivable that the subjects kept their eyes constantly fixated exactly midway between the two lights. This would explain the negative results here obtained. However, questioning of subjects indicated that this type of intermediate fixation did not occur.

their 'central' theory of preparatory set but has expressed grave concern lest "uncritical reading [of this study] will provide Gestalt psychologists with apparent support for their unsubstantiated doctrines of centrally maintained brain dynamics" (p. 627). The results of the present study, together with the collateral evidence which is here cited, seem to constitute a satisfactory refutation of Freeman's hypothesis that all sets necessarily have muscular correlates. (If the more plausible alternative 'central' theory of set provides support for Gestalt views, why is this so lamentable? Is it not the responsibility of scientists to pursue the facts as nature has ordered them and to let their implications fall where they may?)

Elsewhere Freeman (5) has, in fact, taken a position which is tantamount to the acceptance of the 'central' theory of set. He says:

> The major issue in current theorizing about set phenomena is whether or not the covert or incipient muscle activities found to precede reaction differ only in degree from the overt pattern of response. The simplest assumption for our present experiment is that the tension condition of the muscle *per se* rendered it more responsive to external stimulation and that overt reaction was merely a matter of simple accretion in excitation due to such stimulation. This theory may be fairly adequate to simple reaction performance, when a specific stimulus is anticipated by and flows into a specifically prepared channel of response. But when a greater latitude of stimulus possibilities is anticipated, then the interaction of excitation of external origin and that associated with the muscular set requires more complicated assumptions. Here the tonic set cannot be identical with the phasic response, for the overt responses which appear differ in character from the preparatory set. To explain how a given tonic set may sustain a variety of phasic responses we must appeal to the retroflexive mechanisms which places [sic] excitation from the 'set' in the brain field along with the excitation of external origin and make the overt response the *product* of their dynamic interaction. One may argue that such assumptions become necessary even in simple reaction time performances, for the set of the primary reacting muscle would have difficulty in directing the external aroused excitation to itself unless it acted retroflexly to lower the response threshold of the cortical motorneurones concerned with initiating the voluntary efferent discharge (p. 147).

This a well-founded argument. It should be noted, however, that Freeman here concedes that differential preparedness for the performance of specific (phasic) responses does not inevitably involve differential muscular (tonic) backgrounds. Since a given pattern of muscular tension may be common to a variety of different sets, the determining factor must, as Freeman says, be 'the "set" in the brain field.' This statement is different in no essential respect from the hypothesis advanced by Mowrer, Rayman, and Bliss to the effect that preparatory sets are mediated and differentiated principally by neural rather than by neuro-muscular mechanisms, to which, in the paper first cited, Freeman has taken such strenuous exception. Again it should be stressed that this view does not deny that muscular factors are often integrated with neural factors in states of specific response-readiness (to form the kind of dynamic 'product' referred

to by Freeman), but such an acknowledgment is in no way inconsistent with the assumption that the essential determinants in such states are neural.

In view of the fact that sets represent preparatory adjustments to events which have not yet occurred, there is a natural inclination to think of them in teleological terms. This, of course, is unwarranted and unnecessary. Perhaps what Freeman is really inveighing against is the danger of *foresight* being equated to the Gestalt concept of *insight*. Although the forward-looking, anticipatory character of sets is obvious operationally, it must be kept in mind that they always come into being on the basis of some stimulus event that has already occurred. Here the conditioned-response paradigm is an aid to further clarification. A conditioned response, although elicited by a stimulus which necessarily precedes it, is basically anticipatory, providing a preparatory adjustment for the unconditioned stimulus, of which the first (conditioned) stimulus is premonitory. On prior occasions the one stimulus has been followed by the other; by the process of conditioning the affected organism learns, on the basis of *past* experience, to make adaptive adjustments to impending, *future* events. Sets, therefore, may be thought of as *implicit conditioned responses* (10), which, however, may be confined to changes within the central nervous system, without necessarily involving muscular mechanisms. In the experiments reported in the preceding pages, each stimulus served not only to produce an immediate overt response but also acted as a *signal* of the approach of the next succeeding stimulus. In a later paper the question as to how sets become conditioned to stimulus traces will be considered at length. It is only relevant to observe in the present connection that this problem can be satisfactorily handled without departing from the assumption that sets are, or at least may be, purely neural.

Summary

The purpose of the present study has been to obtain further evidence which would bear upon the validity of two rival hypotheses concerning the nature of preparatory set. These hypotheses are:

A. That preparatory set necessarily involves both neural and muscular elements.

B. That preparatory set, as an essentially neural phenomenon, may exist and vary independently of any muscular involvement.

In an earlier study it was found that the quickness with which subjects could make the same, identical response (finger retraction) to two different stimuli varied according to whether the subjects more strongly 'expected' (*i.e.*, were better 'set' for) one type of

stimulus than the other. (Since a given state of muscular tension in the responding bodily member would presumably be equally effective in facilitating reaction to either type of stimulus, the observed differences in promptness of reaction (which were taken to indicate differential sets) were interpreted as due to purely neural, or 'central,' factors.)

This interpretation has been challenged on the grounds that although differential muscular tensions in the responding bodily member admittedly cannot be regarded as responsible in the study cited for the results obtained, there remains the possibility that these results were due, not to purely central factors, but to differential sense-organ adjustments of a muscular nature. In the first three experiments here reported the possibility of differential receptor adjustments has been severely minimized, if not entirely eliminated, by using stimuli which, although perceptively different, affected the same sense modality in such a way that the same receptor adjustment would be adequate for both types of stimuli. Under these conditions the effect referred to above did not appear, *i.e.*, unexpected shifts from one type of stimulus to the other did not result in increased reaction times. This finding was regarded as ambiguous, neither proving nor disproving either Hypothesis *A* or *B*, since the absence of the phenomenon in question might have been due either to the absence of differential sense-organ adjustments or to the absence of cross-modality stimulation.

In a fourth experiment cross-modality stimulation was employed, involving the tactual and auditory senses. The phenomenon in question again appeared, and since there are no known muscles which specifically adjust the tactual receptors for optimal functioning, Hypothesis *A* is regarded as weakened by this finding and Hypothesis *B* strengthened.

In a fifth experiment, two stimuli were again employed in the same (visual) sense modality but they were of such a character as to demand (because of their divergent locations) markedly differentiated receptor adjustments. The failure to obtain the phenomenon in question in this situation again suggests that it is dependent upon cross-modality stimulation rather than upon differential sense-organ adjustments and supports the neural rather than the neuro-muscular theory of set.

Incidental features of the experimental findings are briefly discussed, along with the question of the precise mechanism whereby set, as a purely neural phenomenon, is mediated.

(Manuscript received August 17, 1940)

References

1. Bass, M. J., and Hull, C. L., The irradition of a tactile conditioned reflex in man, *J. comp. Psychol.*, 1934, **17**, 47–65.
2. Dusser de Barenne, J. G. and McCulloch, W. S., Factors for facilitation and extinction in the central nervous system, *J. Neurophysiol.*, 1939, **2**, 319–355.
3. Dusser de Barenne, J. G., and McCulloch, W. S., Suppression of motor response upon stimulation of area 4-s of the cerebral cortex, *Am. J. Physiol.*, 1939, **126**, 482.
4. Freeman, G. L., 'Central' vs. 'peripheral' locus of set; a critique of the Mowrer, Rayman, and Bliss 'demonstration,' *J. exper. Psychol.*, 1940, **26**, 622–628.
5. Freeman, G. L., and Kendall, W. E., The effect upon reaction time of muscular tension induced at various preparatory intervals, *J. exper. Psychol.*, 1940, **27**, 136–148.
6. Hall, G. S., and Kries, J. v., Ueber die Abhängigkeit der Reactionszeiten vom Ort des Reizes, *Arch. f. Physiol.* (*Supplement-Bund*), 1879, 1–10.
7. Lorente de Nó, R., Vestibulo-ocular reflex arc, *Arch. Neurol. Psychiat.*, 1933, **30**, 245–291.
8. Mowrer, O. H., The electrical response of the vestibular nerve during adequate stimulation, *Science*, 1935, **81**, 180–181.
9. Mowrer, O. H., The modification of vestibular nystagmus by means of repeated elicitation, *Comp. Psychol. Monogr.*, 1934, **9**, pp. 48.
10. Mowrer, O. H., Preparatory set (expectancy)—some methods of measurement, *Psychol. Monogr.*, 1940, **52**, pp. 43.
11. Mowrer, O. H., Rayman, N. Nelson, and Bliss, Eugene L., Preparatory set (expectancy) —An experimental demonstration of its 'central' locus, *J. exper. Psychol.*, 1940, **26**, 357–372.
12. Pavlov, I. P., *Conditioned Reflexes*. London: Oxford University Press, 1928, pp. 430.
13. Winer, P. and Mowrer, O. N. Further experiments on the locus of preparatory set. (To be published.)

THE CONDITIONING OF THE BLOCKING OF THE ALPHA RHYTHM OF THE HUMAN ELECTROENCEPHALOGRAM [1]

BY JOHN R. KNOTT

Departments of Psychology and Neurology, University of Iowa

AND

CHARLES E. HENRY

Bradley Home and Brown University

The alpha rhythm of the electroencephalogram (EEG) is in most individuals a relatively transitory phenomenon. That is, it 'blocks' from time to time even though no known stimuli are introduced into the environment of the organism. Experimentally, this blocking can be produced by a wide variety of sensory stimuli (**1**, **2**, **6**, **7**, **8**, **9**, **10**, **11**) and it has been found that the spontaneous blocking may be related to certain 'conscious' processes not associated with detectable stimuli (**12**).

The understanding of the variability of the electroencephalographic pattern, and especially the variability of the alpha rhythm, is one of the key problems in the interpretation of the psychological significance of the EEG (**5**), and the experimental production of a variability which is apart in time from the presentation of a particular sensory stimulus should afford a fairly direct attack upon it. Further, the experimental production of a variability in a cortical process (*i.e.*, the blocking of the alpha, a cortical response) which would come under the general heading of learning might provide data which could serve as an orientation for the further investigation of cortical responses as related to the learning process.

The production of a 'conditioned' variability of the EEG has been previously approached by several workers. Loomis, Harvey and Hobart (**6**) made note of the fact that an auditory stimulus, which was relatively impotent in the production of alpha blocking, when presented simultaneously with a visual stimulus, which was relatively potent in the production of alpha blocking, would, after several paired presentations, produce blocking in the absence of the visual stimulus. Cruikshank (**2**) observed that when light was to serve as

[1] The present report is one of a series of papers under the direction of the senior author and dealing with physiological indexes of psychological phenomena. All of the data were collected and evaluated in the laboratory of physiological psychology, in the Department of Psychology, University of Iowa. This report was presented at the meetings of the Midwestern Psychological Association, May 3, 1940, in Chicago, Ill.

a signal for a manual response, an auditory preparatory stimulus presented at a constant interval before the light would elicit blocking prior to the appearance of the light. Travis and Egan (**10**), attempting to study quantitatively the effect which the previous workers had merely noted, studied the response of the alpha rhythm to an auditory stimulus before, during and after pairing it with a visual stimulus. The total duration of their auditory stimulus was .75 sec., the total duration of their visual stimulus was the last .25 sec. of this interval. Their data may be summarized as follows: tone prior to light was effective in blocking the alpha rhythm in 11 percent of the stimulus intervals; tone paired with light was effective in 34 percent of the stimulus intervals; after pairing with light, tone was decreasingly effective, and at the end of 3 or 4 such single presentations was no more effective than prior to pairing.

In evaluating the data of Travis and Egan we were interested in the fact that tone was just as effective in producing alpha blocking in the first one-fifth of the paired series as in the subsequent similar divisions. That is, there was not apparent in the presentation of their data a curve of increasing effectiveness as pairing was continued; the increment seemed to be initial and abrupt. This does not look like conditioning, but rather like sensitization ('alpha' conditioning).

We were also interested in the relationship of this abrupt increment to data collected in this laboratory by Dr. B. Martinson (unpublished). Martinson apparently reproduced with no pairing whatsoever the response of the alpha rhythm to tone which Travis and Egan reported during pairing. The technique used by Martinson was as follows: two tones of equal frequency but of differing intensities, and two lights of differing intensities were presented in random order. The tone (*i.e.*, all responses to tone, regardless of intensity) produced blocking in 42 percent of the stimulus intervals. This is slightly higher than the 'conditioned' effectiveness reported by Travis and Egan.

These data apparently indicate a 'sensitization' effect which is not to be confused with 'conditioning' in the usual sense of the term. The fact that Bakes (**1**) observed an increase in the effectiveness of tone when tone was a signal for manual response, would perhaps suggest that a peripheral factor involved in the set for response is important in the sensitization.[2]

Since Travis and Egan did not report increasing effectiveness and since a sensitizing of the response to tone could not be ruled out as a factor influencing their data, it was believed necessary to demonstrate the former and experimentally delimit the latter before it would be

[2] The peripheral factor in the Bakes experiment is inferred from control experiments introduced by Knott (**4**) in research with visual stimuli as signal for response.

possible to erect hypotheses concerning the nature of the variability of the alpha rhythm or concerning the role of the cortex in learning.

The present series of experiments was undertaken to discover whether or not these factors could be studied and 'conditioning' of the blocking of the alpha verified; and if such an effect could be verified, to point out some of the implications of the data.

Apparatus and Procedure

The apparatus used for recording the EEG consisted of high-gain, low-frequency amplifiers and a Grass multi-element ink-writing oscillograph. Monopolar leads were taken from the left occipital, motor and frontal regions. The data to be presented pertain only to the left occipital area. It may be noted that the three areas recorded did not respond in an identical manner and that the occipital lead data cannot be used to describe the action of the electrical activity of the whole cortex.[3]

The stimuli used were tone (conditioned stimulus) and light (unconditioned stimulus). The tone (180 cycles) was adjusted in intensity until it was judged by the subject to be near threshold. (Precise psychophysical techniques were not used for these determinations.) This was done to minimize the effect of the tone *per se* upon the alpha rhythm. The light was a 150 watt frosted bulb which was enclosed behind a 3 × 5 inch ground glass screen.

The subjects were seated in a dark, sound proofed room. The tone was presented through a loud speaker above their heads; the light was three feet in front of them. The subjects were instructed to keep their eyes open.

The durations of the stimuli and their temporal relationships were controlled by means of a photoelectric timing device especially constructed for this purpose.[4] The principle of this timer is as follows: two independent photoelectric relay units are activated by a single light source which transmits its energy through adjustable sectors in two discs attached to a common axle. The discs rotate at a constant speed, being driven by a synchronous motor. By controlling the width of the sectors, the durations of the stimuli are adjustable over a fairly wide and continuous range; by controlling the relationship of the sectors to each other, the intervals between stimuli are adjustable through a like range. A manually operated clutch built into the driving mechanism makes it possible to control the intervals between successive stimulus presentations. The variability of the system, checked with a Springfield Electric Timer, was found to be less than 1 percent when the line voltage was held constant by manual control.

Inasmuch as the problem involved a study of a conditioned response which seemed likely to become confused with a sensitized response, it appeared to be necessary to separate the occurrence of the two responses in time. For this reason it was arbitrarily decided to investigate delayed responses in which the interval between the conditioned and unconditioned stimuli was of sufficient magnitude to make discrimination of sensitization and anticipatory conditioned responses possible; since the recovery of the alpha rhythm subsequent to blocking induced by auditory stimulation is quite rapid, this technique seemed suitable.

In a series of exploratory experiments, delayed responses with delay intervals as long as 9 sec. were studied; others were studied with delay intervals as brief as 2 sec.; several short trace experiments were also attempted, but with results which were somewhat equivocal, although suggestive. On the basis of this series of preliminary experiments, it was decided to use an interval of 4 sec. between the conditioned and unconditioned stimuli. Tone was thus presented for a total of 5 sec., and light was presented during the last 1 of these.

The preliminary observations had led us to believe that the 'conditioned' response to be studied was very labile and subject to a great deal of spontaneous variability. Therefore, it was decided to present 75 paired stimulations to each subject. The tone, which was originally more effective than tone is generally reported to be, was presented 25 times in an adaptation series;

[3] Research is now under way to determine whether or not systematic variations in these three areas do occur, or whether the variations noted are to be described as 'chance.'

[4] The authors wish to acknowledge the assistance of Mr. P. E. Griffith in the design and construction of this instrument.

after the pairing it was again presented 25 times in the extinction series. Six subjects, selected on the basis of a strong and relatively unvarying alpha rhythm, were studied. Since the problem involved the production of variability it seemed more justifiable to study a fairly constant rhythm than to study an initially variable rhythm.

Analysis of Data

The method of analysis involved an all-or-nothing rating of 'blocking' of the alpha within certain arbitrary segments of the record. In order to prevent subjective bias from influencing the evaluation of the effect of the stimuli on the EEG, blocking of alpha was defined as a 50 percent reduction in alpha amplitude (relative to the alpha amplitude at the beginning of the interval studied) in 3 or more alpha waves.

The total stimulus interval was divided into 4 sub-intervals and this criterion was applied to each. The first sub-interval was the one second following the initiation of tone; the second sub-interval was the *next* one and one-half seconds; the third sub-interval was the *next* one and one-half seconds; and the last sub-interval was the final one second of the total interval (the second during which light appeared during the paired trials). Blocking, as defined, was measured in these intervals before pairing, during pairing and after pairing. The three experimental periods are referred to as TBL, TWL and TAL in the graphic presentation which follows.

In the text and figures the four subintervals are referred to as E_T, A_1, A_2, and E_L, respectively.

Since the ratings of blocking are all-or-none in character, the treatment of the data involved a grouping of trials and the computation of the percentage of that grouping in which the criterion of blocking was met. Changes in percentage effectiveness (in blocking) were computed by means of the formula giving the standard deviation of a percentage and the usual formula for the ratio of the difference to the standard error of the difference. This particular technique is somewhat inadequate, but there is no other which can be used.

Results

The data from the six subjects are presented graphically in Fig. 1*a*, 1*b*, 1*c* and 1*d*, and tabularly in Table 1. In Fig. 1*c* and 1*d*, the trials are grouped by 5's to show the changes in rating on a somewhat

TABLE 1

Percentage of Intervals in which the Criterion of 'Blocking' was Met

Trials	Intervals							
	E_T		A_1		A_2		E_L	
	N	%+	*N*	%+	*N*	%+	*N*	%+
TBL 1–10....	48	58	46	41	45	31	45	25
TBL 16–25....	44	23	45	33	45	33	43	30
TWL 1–10....	44	39	46	41	44	43	44	59
TWL 11–20....	41	37	41	39	41	54	41	63
TWL 21–30....	45	51	45	49	45	58	45	60
TWL 31–40....	31	35	31	31	31	39	31	39
TWL 41–50....	33	48	33	61	31	29	33	58
TWL 51–60....	36	32	39	31	36	37	36	50
TWL 61–70....	30	23	30	30	32	43	30	63
TWL (71–75)...	(14)	(29)	(14)	(57)	(14)	(57)	(14)	(43)
TAL 1–10....	38	63	37	46	38	55	38	34
TAL 16–25....	35	40	33	33	33	52	33	21

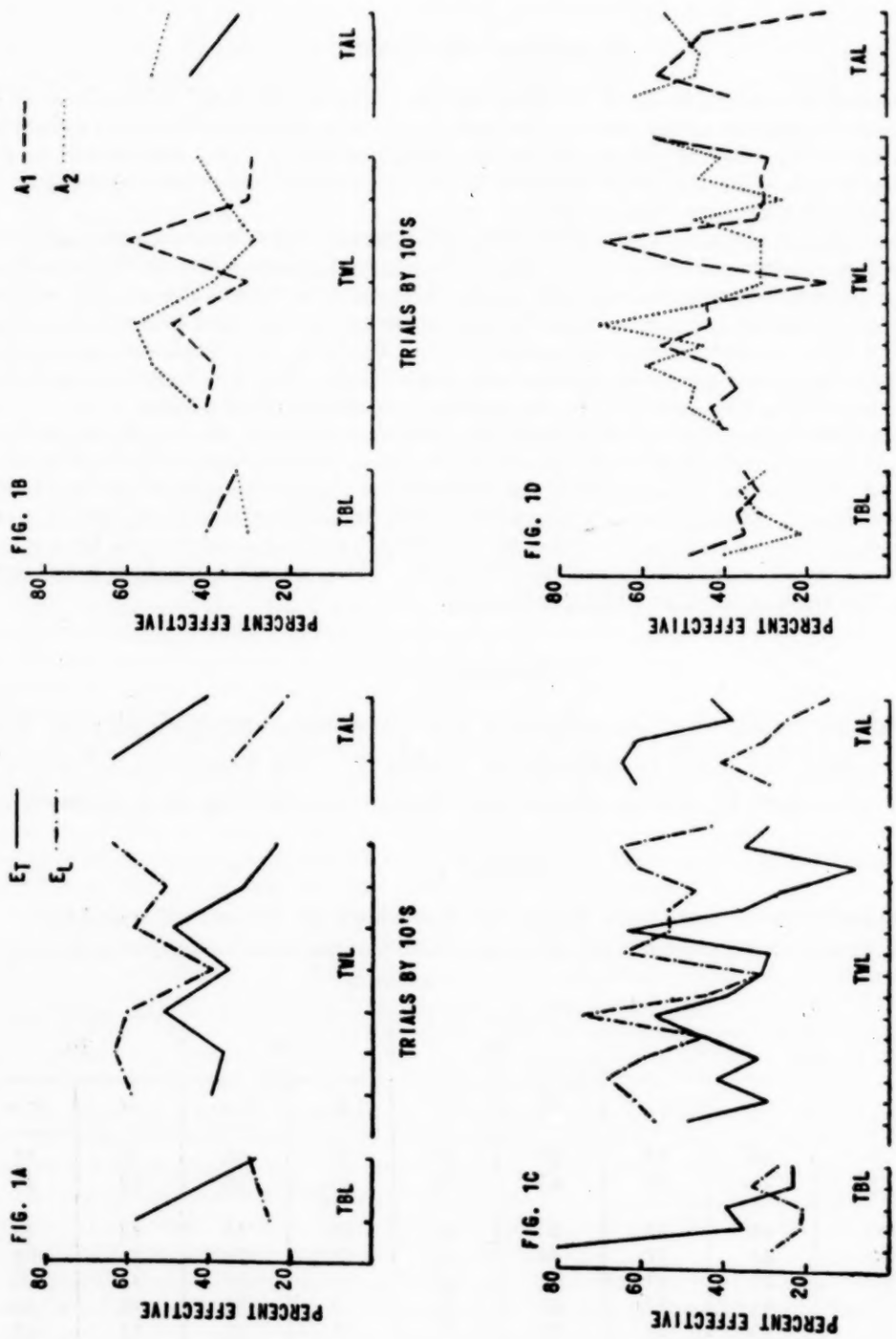


FIG. 1. Adaptation, conditioning and extinction curves for the time intervals studied. For further description see text.

wider time-scale than in Fig. 1*a* and 1*b*, in which the trials are grouped by 10's. This latter grouping is the one which was used for the statistical treatment, since larger N's were thus obtainable (when grouped by 5's, the N's were frequently so small as to make application of the formulae undesirable). The last 5 trials of the paired series were thus dropped from consideration in this grouping.

Before Pairing with Light.—During the 25 adaptation presentations of tone, interval E_T presented a sharp decline in effectiveness. Comparing the first and last ten trials, the drop was statistically significant (critical ratio = 3.7).[5] The variations in intervals A_1, A_2 and E_L were not significant.

During Pairing with Light.—Certain changes in the blocking characteristics of the intervals studied were noted when light was introduced into the total stimulus interval. There was an immediate increase in the effectiveness of tone (E_T) which was not significant. In trials 21–30 E_T reached a peak of effectiveness which was significantly greater than the value of the last ten trials of the adaptation series (critical ratio = 2.9). The subsequent sharp decline in E_T in the next ten trials was not significant; the lowest value reached during pairing, in trials 61–70, was marginally significantly lower than the peak value (critical ratio = 2.6).

Interval A_1 tended to follow the same course of events as did interval E_T, probably due to their close proximity in time. The peak value of A_1 was reached in trials 41–50, and this was marginally significantly greater than the value of the last ten adaptation trials (critical ratio = 2.5). Within the next ten trials, the drop was significant (critical ratio = 2.7).

Interval A_2, which lay just prior to the appearance of light in the paired series, increased in effectiveness within the first 10 trials of paired presentations of tone and light, although this increase above the last ten adaptation trials was not significant. Even the peak value, in trials 21–30, was only marginally greater than the last ten adaptation trials (critical ratio = 2.5). The decline in effectiveness from this time on was not significant when successive groups were compared; when the lowest value (trials 41–50) was compared with the highest value (trials 21–30), the decline was marginally significant (critical ratio = 2.6). The subsequent rise was not significant.

The effectiveness of E_L rose immediately after the presentation of light. It is to be noted that the value reached was not 100 percent. This is due to the fact that the criterion of blocking was not always

[5] 'Significance' was arbitrarily set at a ratio of 2.7; ratios between 2.2 and 2.7 were regarded as 'marginally' significant.

reached.[6] The increase in the percentage of blocking occurring in the first ten actual light stimulations as compared with the same interval when no light appeared (the last ten adaptation trials) was statistically significant (critical ratio = 2.8). The decline noted in trials 31–40 was not significant; nor was the subsequent rise.

After Pairing with Light.—The percentage of blocking occurring in intervals E_T, A_1, and E_L declined systematically, if not significantly, in the extinction series. Interval A_2 decreased insignificantly.

When Fig. 1*a* is examined, it is seen that there was an increase in the effectiveness of the E_T, A_1 and E_L intervals within the first ten trials, with a subsequent decline.

Comparing the extinction series values with the paired series values, it was found that the first ten extinction trials were more effective than the last ten paired trials in interval E_T (critical ratio = 3.6). The increase in A_1 was not significant. The decrease in E_L was considered marginally significant (critical ratio = 2.5). In this last interval, the decrease from paired trials 61–70 to extinction trials 16–25 was significant (critical ratio = 3.7).

When effectiveness values for the adaptation series, trials 16–25, were compared with the effectiveness values for trials 1–10 of the extinction series, E_T increased significantly (critical ratio = 4.0), while A_1, A_2 and E_L did not change significantly.

Discussion

The data just presented seem to indicate that continued pairing of a 'conditioned stimulus' and 'unconditioned stimulus' did not bring about continued increases in the response to the 'conditioned stimulus.' However, for a part of the series of paired presentations, that is, during the first 30 trials, increments in one area of the records (A_2) did occur, and these appeared to be systematic. Had the pairing been discontinued at this point, one would have been led to the conclusion that conditioning of the blocking of the alpha rhythm had occurred.

It is known that continued reinforcement leads to a type of inhibition ('inhibition of reinforcement') (3). When this phenome-

[6] The authors would like to state that there can be no admission of data which are not defined in some rigorous terms. The fact that light has been regarded to be 'highly effective' in blocking the alpha rhythm, and that tone has been regarded as 'relatively ineffective' in blocking the alpha rhythm may lead to an inaccurate registering of the data. For instance, when it is known that light has been presented, an experimenter may be led to call a slight shift in the amplitudes of alpha an alpha block; the same quantitative disturbance subsequent to the presentation of an auditory stimulus may be used as an example of the relative impotency of tone in producing alpha blocks. It is for this reason that a definite, rigorous criterion of blocking was herein established. It is probable, however, that a *continuous* measure of alpha amplitudes would be the best method of handling the data.

non of a decrement in response with increased reinforcement occurs, cessation of reinforcement is associated with an increment in the response to the conditioned stimulus. This appears to be the phenomenon which has occurred in the present series of experiments, for all of the intervals subjected to analysis exhibited an initial increase in value after reinforcement had ceased. When the data were plotted by 5's (Fig. 1*a*) this was particularly clear.

That some phenomenon which could be called 'inhibition' developed during the period of reinforcement seems to be obvious from an inspection of the data. Increments in the E_T, A_1 and A_2 values appeared during the first 30 paired trials, but within the next ten trials there was a decrement in response in all values, including those of E_L, the supposedly potent unconditioned stimulus. This inhibition seems, therefore, to be a generalized process. It is to be noted further that there were fewer available alpha waves to be blocked; the N's drop in magnitude also. (It is interesting that a decrease in alpha is known to be associated with the earliest stages of sleep. A period of 'inhibition' may be either a condition of sleep, *per se*, or a condition resembling sleep, on the basis of the present data. Wenger (13) has made observations which would support the former view. The subjects reported a feeling of drowsiness in the middle portions of the experiment.)

Following this generalized period of response decrement, there appeared a temporary increase in E_T and A_1, and a permanent increase in E_L. Due to the latter, it would be difficult to conceive of the subsequent decrement in E_T and A_1 as being associated with a process of generalized inhibition. The decrement in the E_T and A_1 values, might, however, be associated with a process similar to the 'adaptation' process noted in E_T during the pre-pairing stimulation. If so, the adaptation was specific to the experimental situation of pairing, for it no longer held when pairing was stopped. E_T then underwent an immediate increase with, subsequently, further adaptation.

The mechanism of the observed adaptation is exceptionally difficult to evaluate. One view which could be taken would attribute the decrease in the effectiveness of tone to an inhibitory process. The change in the situation associated with the presentation of tone plus light could be interpreted as bringing about dis-inhibition, and during the subsequent presentation of the paired stimuli further adaptation would be expected to take place. Cessation of pairing, again introducing a shift in the situation, could be interpreted as a further disinhibitor, with further adaptation (inhibition) accompanying repeated stimulation.

The gradual decline in the E_T values during pairing tends to fit this latter evaluation. It should be noted that this is the interval

which was studied by Travis and Egan. Their data, therefore, may be considered in terms of such an evaluation as well as in terms of a conditioning interpretation. This would thus amount to sensitization of the original but adapted response to tone. Whether or not disinhibition is another name for the same phenomenon, is not clear.

The interval A_2, that which fell just prior in time to the light stimulus, exhibited in the present series of experiments variations which are somewhat more consistent with data obtained in conditioned responses less labile than the blocking of the alpha rhythm of the EEG. The blocking in this interval can be characterized as 'anticipatory.' It may be significant that as the A_2 values declined following trials 31–40 (which may have been associated with a generalized inhibitory process), the values of interval A_1 increased. This can be interpreted as representative of the tendency of the anticipatory response to move closer to the conditioned stimulus as pairing continues. During trials 41–50 the value of A_1 was significantly greater than the value of A_2 (critical ratio 2.7). The shift in subsequent trials appeared to be essentially inverse, and opposite to the direction begun in trials 41–50.

Any conclusion to the effect that the blocking of the alpha rhythm is a response which can be conditioned seems to rest upon observations of a labile phenomenon. The authors would, however, offer the tentative conclusion that conditioned anticipatory blocking of the alpha rhythm appears to be possible, that it may be masked by inherently spontaneous fluctuations in the response being studied, and that, due to its apparent susceptibility to inhibition of reinforcement, it can better be investigated by fewer paired trials.

If this tentative conclusion be substantiated by further work, it may be possible to make certain extensions in psychophysiological theory. One such extension would be in a direction away from the so-called motor theories of psychological processes and those modifications of them which postulate an association between what must be fairly precise movement-produced stimuli (conditioned stimuli which have their origin in not-too-precisely observed implicit movements which are essentially inferred from macroscopic investigations) and an unconditioned response. Instead, it may be possible to postulate an association between a cortical state following a peripheral stimulus (conditioned stimulus) and the cortical state preceding the unconditioned response. This view would not deny movement-produced stimuli in the total response situation; it would, instead lead to the investigation of their relative role in the situation and their relation to cortical mechanisms of the type just indicated.

Relative to the significance of the conclusion (that blocking of the alpha rhythm may be conditioned) for possible psychological

interpretations of the EEG, an essentially identical postulation would be made. It would, in this light, appear that the spontaneous fluctuation of the alpha rhythm could represent a series of intra-cerebral 'stimulus'-'response' relationships. A peripheral stimulus could set off such a series, and this series could continue beyond the temporal duration of the peripheral stimulus or its specific cortical response. Experimental verification of the simpler systems of this type can be made.

Summary and Conclusions

1. The conditioning of the blocking of the alpha rhythm of the human EEG has been investigated. Previous work on this problem appears not to have established sufficient experimental controls to enable differentiation between sensitization of the previously adapted response to the conditioned stimulus and the response believed to be the conditioned response. To make such differentiation feasible, the conditioned and the unconditioned stimuli were separated in time by 4 sec., in order that full recovery of the alpha rhythm could occur after its response to the conditioned stimulus and before the unconditioned stimulus was presented.

2. A conditioned anticipatory response which was unstable and appeared subject to inhibition of reinforcement appears to have been established. Sensitization of the response to the conditioned stimulus appears to have occurred, and this serves to explain the unorthodox type of conditioning data reported by previous workers.

3. The implications of the conclusion that conditioning of the alpha rhythm can be accepted are discussed, with full allowance for the tentative form of the statement. These implications include the postulation of intra-cerebral association mechanisms which are basically independent of movement-produced stimuli but which could operate in conjunction with them. Further work involving a study of the relationships between these physiological mechanisms is necessary.

References

1. Bakes, F. P., Effects of response to auditory stimulation on latent time of blocking of the Berger rhythm. *J. exp. Psychol.*, 1939, **24**, 406–418.
2. Cruikshank, R. M., Human occipital brain potentials as affected by intensity-duration variables of visual stimulation. *J. exp. Psychol.*, 1937, **21**, 625–641.
3. Hovland, C. I., 'Inhibition of reinforcement' and phenomena of experimental extinction. *Proc. Nat. Acad. Sci.*, Wash., 1936, **22**, 430–433.
4. Knott, J. R., Some effects of mental set upon electrophysiological processes of the human cerebral cortex. *J. exp. Psychol.*, 1939, **24**, 384–405.
5. Knott, J. R., Electroencephalography and physiological psychology: Evaluation and statement of problem. (To be published.)

6. Loomis, A. L., Harvey, E. N., and Hobart, G. A., Cerebral states during sleep as studied by human brain potentials. *J. exp. Psychol.*, 1937, **21**, 127–144.
7. Travis, L. E., Knott, J. R., and Griffith, P. E., Effect of response upon latency and frequency of the Berger rhythm. *J. gen. Psychol.*, 1937, **16**, 391–401.
8. Travis, L. E., and Knott, J. R., Brain potential studies of perseveration: I. Perseveration time to light. *J. Psychol.*, 1936, **3**, 97–100.
9. Travis, L. E., and Knott, J. R. Brain potential studies of perseveration: II. Perseveration time to visually presented words. *J. exp. Psychol.*, 1937, **21**, 353–358.
10. Travis, L. E., and Egan, J. P., Conditioning of the electrical response of the cortex. *J. exp. Psychol.*, 1938, **22**, 524–531.
11. Travis, L. E., and Barber, V., The effect of tactile stimulation upon the Berger rhythm. *J. exp. Psychol.*, 1938, **22**, 269–272.
12. Travis, L. E., Brain potentials and the temporal course of consciousness. *J. exp. Psychol.*, 1937, **21**, 302–309.
13. Wenger, M. A., A criticism of Pavlov's concept of internal inhibition. *Psychol. Rev.*, 1937, **44**, 297–312.

(Manuscript received December 2, 1940)

FACTORS WHICH WORK TOWARD UNITY OR COHERENCE IN VISUAL DESIGN

BY IVY G. CAMPBELL

Wells College

Perhaps the most universally demanded characteristic of a good design is that it have unity. Sometimes it is added that there must be unity in variety. But when one seeks to find just what is meant by these demands he finds no one definition or meaning agreed upon and often indeed he finds ambiguity or indefiniteness of meaning. In the Oxford dictionary definition of unity as 'oneness, being one or single or individual, being formed of parts that constitute a whole, due interconnexion and coherence of parts,' we have at least three clearly separable meanings of unity indicated: (1) a one entity or thing with no consideration of whether or not made up of parts; (2) a whole formed of parts; (3) a whole of coherent parts. Now the question arises, is a good design one in which its unity is such that no sense of separateness of parts exists, or one in which variety of parts is felt to make up the whole, or one where coherence of parts in a whole is felt.

Clearly a number of psychological problems are involved in the above question. When we see a design: (1) do we or do we not see it as made up of parts? (2) Are there particular characteristics of a part which determine whether or not it shall be seen as a part? (3) Does the degree of coherence of the parts have an effect upon the relative ease with which we see the parts? (4) What makes for coherence and are there certain characteristics which belong to parts that cohere? (5) Do these same characteristics belong to the parts of a design which lacks coherence?

The following experiment was performed in 1931 at Wells College in collaboration with Marianna Platt to throw light upon some of the points in the above problems, particularly to answer the question: does the degree of unity or coherence of a design correlate positively with the degree of perceptual loss of separateness of parts, or rather, do characteristics of the parts, whether parts of coherent or not coherent designs, determine the loss of separateness of parts?

Designs which we considered unified (what we shall call good) and those not so unified (poor) were constructed.[1] Good designs were such as we considered exemplified the æsthetic principles of balance, proportionate division, rhythm, tonality, harmony, nodality —the six characteristics the writer believes present in all good designs.

[1] Constructed by Miss Platt.

Poor designs were such as we thought lacked these principles. By tonality is meant a consistency in the types of lines or shapes used so that a resultant feeling of oneness of tone emerges. Harmony means that the types of lines or shapes used are consonant with one another. Nodality means that there are certain positions in space, time, or action that possess greater importance than do others. In visual designs this principle is regarded when divisions or emphases are made at such important positions. Rhythm means repetition used structurally. The usual meanings of balance and of proportionate division were accepted. As will be seen in the report which follows, the final classification of the members of the series into good and poor

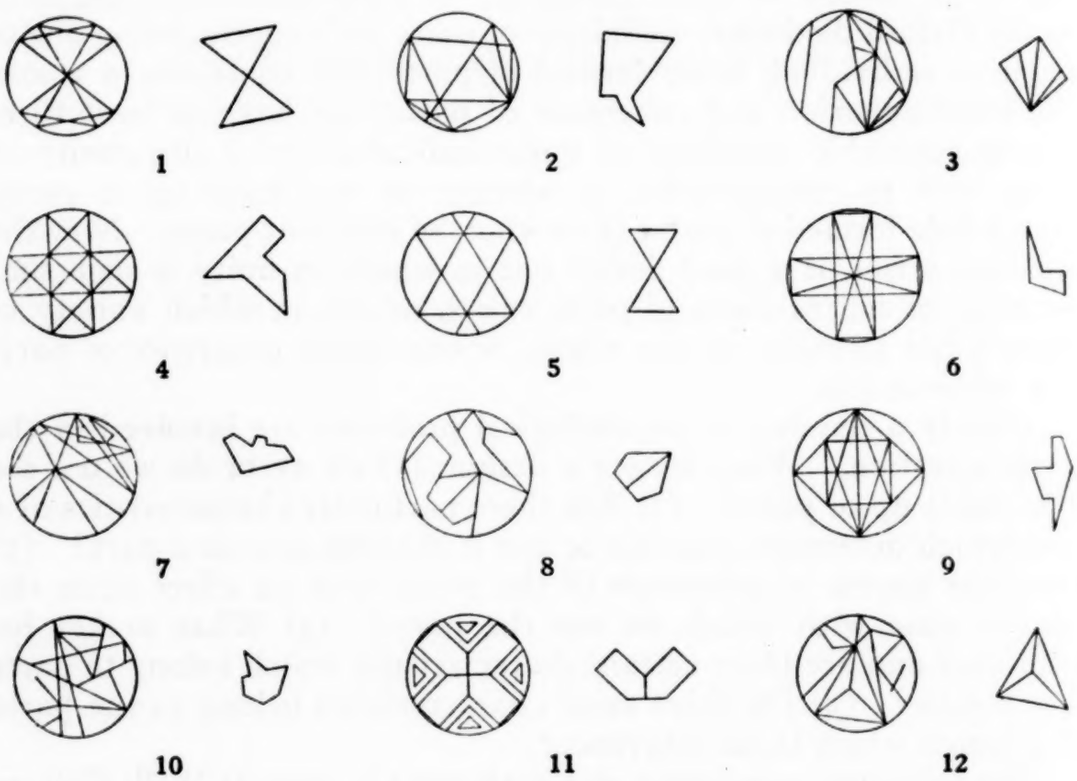


PLATE I

designs was based upon the agreements of the Os who took part in the experiment. It should be added that we did not consider our good designs as excellent artistically but only that they exemplified the principles structural to such designs.

For the first series of observations twelve designs, six of which we considered good, six poor, were constructed. All designs consisted of straight lines composed within a circle of $2\frac{1}{2}''$ diameter and each was drawn in black India ink on the left side of a card $4'' \times 8''$. On the right side of the card was drawn a part of the whole design shown on the left. In the case of six of the cards, three of good de-

signs, three of poor, the part drawn was considered to be a 'natural' part of the larger figure. In the other six cards, three good and three poor designs, the part was considered 'unnatural.' Just what the factors are that make for naturalness and unnaturalness of part will be discussed later. Here it will only be stated that a natural part seemed one which was perceptually a structural unit of the whole design, while an unnatural one, although physically as truly a part of the whole design as was the natural, yet did not seem such perceptually.

From the five parts of the experiment the following data were secured: (1) Os' judgments as to the relative goodness of the different designs, (2) as to the naturalness and unnaturalness of parts; (3) the times it took to see the part figures in the whole designs; (4) the parts of the designs which were most dominant to perception.

Part I

15 Os

Presentation 1

Here the whole designs were shown one at a time in the Gottschaldt exposure apparatus. This is an apparatus in which the opening in an episcotister disc automatically and gradually increases and decreases in size, so that a source of light placed behind the disc and viewed through it becomes constantly stronger, and then weaker. When placed between a projecting lantern and a screen, at first, when the disc has no opening, nothing is seen on the screen; gradually, with the increase in the size of the opening, the screen image becomes ever and ever clearer until full visibility is reached—then gradually it fades from vision.

The Os were instructed to draw the first and the last things they saw as the figures slowly came into and then slowly faded from vision. The experimenters hoped by this method to obtain drawings of the natural parts of the large figures. The writer had already used this method in studying the composition of works of art (designs, pictures, statues) and had found that Os drew the skeleton, structural design. She argued, therefore, that if a part were structural and hence natural this should be indicated by the drawing.

In the drawings of the designs with natural parts, a large majority of the Os drew those parts which the experimenters had judged natural in the cases of designs 1 and 11, and a smaller majority in design 3, but drew other parts for designs 5, 8, 12. For 5 and 12 there was large agreement in the parts drawn, while for 8 no such agreement. In the drawings of designs with unnatural parts none of the Os drew figures even similar to the parts judged unnatural by the experimenters nor did they agree in the parts they did draw. If the assumption is correct that had these been natural parts they would have been drawn, then since they were not drawn they can be assumed to be unnatural. The results of this presentation were utilized in making up the series used in *PART II*.

Presentation 2

The designs were projected one at a time on a daylight screen and the Os were asked to make immediate judgments as to whether they were good, medium or poor. No standards or criteria for goodness were given. Spontaneous, not consciously reflective, judgments were desired. Except in the cases of designs 1 and 3 there was large agreement (70–100 percent) with the experimenters.

Presentation 3

The designs and the part figures at the right were projected one at a time on the screen. The Os were instructed to 'see out' the part figure in the whole design. They were warned not just

to find or trace the small figure, but to see it as a well-defined unit. By means of a stop-watch the time that it took each O to see out each part was recorded. Only whole seconds were recorded. At the end of two minutes the images were removed from the screen, note being taken of the Os who had not 'seen out' the part by this time. The aim of this procedure was: 1) to determine whether natural parts are seen out more easily (judged by the time it takes) than are unnatural; in other words, do natural parts tend to break up the perception of a design into parts or, on the other hand, to make for stronger integration of design; 2) to determine in which designs—good or bad—it is easier to see out parts; to learn, in other words, whether good or bad designs break up more easily into parts.

Results of Part I showed that natural parts are in all cases seen out more easily than unnatural, and that goodness or badness of design has comparatively little influence on the ease with which in perception a design is broken up into parts. The following averages show this clearly. The average time for the six good designs was 13.4″: six poor, 21.5″: six natural, 2.4″: six unnatural, 32.6″: three good natural, 2.0″: three good unnatural, 25.0″: three poor natural, 2.9″: three poor unnatural, 40.2″. Thus in both the good and the poor designs it took over twelve times as long to see out an unnatural part. Both in the cases of natural and of unnatural parts those of good designs were seen out somewhat more quickly than those of poor designs.

Part II

In this part of the experiment the series of *PART I* was used with the following exceptions In *Presentation 1*, other designs were substituted for 3 and 8 and the parts of designs 5 and 12 were drawn to accord with the drawings of the Os in Part I, Presentation 1. In *Presentation 2*, a new design, 13, was added to the series of *Presentation 1*. In *Presentation 3*, designs 3, 8, 9, 13, 14 (a new design) were shown. For *Presentation 1* there were twenty-five Os not used in Part I; for *Presentation 2*, nine Os not used before; for *Presentation 3*, twenty-three of the Os of *Presentation 1* repeated their observations on designs 3, 8, 9, and made observations on designs 13 and 14. In each presentation the Os were instructed to judge whether the part figure was a natural or an unnatural part of the whole design.

The results were as follows. Os' judgments in regard to designs 8, 9, 14 agreed with the experimenters' judgments only 48, 41, and 31 percent respectively; design 3, 58 percent; design 12, 61 percent; designs 7 and 10, 67 percent; designs 4 and 6, 70 percent; design 13, 73 percent; design 11, 76 percent; designs 1, 2, 5, 85 percent.

Part III

On the basis of the results secured in *PARTS I* and *II* a series of designs was made up for this part of the experiment and presented to eighteen new Os. This series is shown in Plate I. All of the designs (and with the same numbering) except 8 and 9 were from those used in the three presentations of *PART II;* design 14 of *PART II, Presentation 3*, here becomes design 8; 9 is a new design; 13 was omitted. In *Presentation 1* of this part the procedure and instructions were the same as in *PART I, Presentation 2;* in *Presentation 2*,

those of *PART I, Presentation 3;* in *Presentation 3*, those of *PART II.* Table I gives the results of the three presentations.

Table I shows: (1) that the Os agreed as to the goodness and poorness of the designs with the plan of the series. For only two of the good designs did as many as one O judge them poor and for only four of the poor as many as one O judge them good. There were a number of medium judgments, but rarely did these approach in number those of good and poor, and in no case did they outnumber them; (2) that the natural parts were seen out much more quickly than the unnatural; (3) that the Os largely agreed in regard to all of the designs as to the naturalness and unnaturalness of parts with the

TABLE I

CLASSIFICATION ACCORDING TO EXPERIMENTERS; Os' Js OF GOOD, MEDIUM, POOR; AVERAGE AND MEDIAN TIME TO "SEE OUT" PART IN WHOLE; Os' Js OF NATURAL AND UNNATURAL PARTS

18 Os

Design	1	8	5	12	3	11	4	6	2	7	10	9
Class	GN	PN	GN	PN	PN	GN	GUn	GUn	PUn	PUn	PUn	GUn
No. Good Js	13	1	14	0	1	17	16	10	1	1	0	12
No. Medium Js	4	7	4	2	3	1	2	7	5	0	1	6
No. Poor Js	1	10	0	16	14	0	0	1	12	17	17	0
Av. Time	1.38″	1.77″	1.94″	2.11″	2.77″	2.83″	19.00″	20.72″	24.66″	30.44″	47.33″	56.11″
Median	1″	1″	1″	1″	3″	2″	16.50″	17.50″	12″	17″	20″	22″
Os not see *	0	0	0	0	0	0	0	0	0	3	3	6
No. Js Nat.	13	11	16	15	11	14	2	2	1	1	2	2
No. Js Unnat.	5	7	2	3	7	4	16	16	17	17	16	16

* Os who at end of 2′ had not seen the part. The time for these Os was counted as 120″.

plan of the series; except for designs 3 and 8 the agreement was over 80 percent. These two designs were planned as poor with natural parts and the fact that so many Os judged them as unnatural suggests that the very conception of naturalness of part of a poor design is perhaps ambiguous. This point will be discussed later.

The data of *PART III* were then worked up to show the average time taken to see out the part for several classifications of the designs. The number in the parenthesis indicates the total number of times that Os failed to see out the part in the two minutes allowed. Average time for the six good designs 16.99″ (6); six poor designs 18.34″ (6); six designs with natural parts 2.13″ (0); six designs with unnatural parts 33.21″ (12); three good designs with natural parts 2.05″ (0); three good designs with unnatural parts 31.94″ (6); three poor designs with natural parts 2.21″ (0); three poor designs with

unnatural parts 34.47″ (6). These data show very strikingly that a natural part of a design is seen out more easily than is an unnatural part and that this is true of both good and poor designs. They also show that the natural and the unnatural parts of a good design are seen out only a little more quickly than are these same parts of a poor design.

Part IV

In 1937 the series of designs used in *PART III* was presented to a group of twenty-one new Os with the same procedure and instructions as in *PART III*, except that judgments of good and poor instead of good, medium and poor were called for. The Os' judgments on naturalness and unnaturalness of part agreed even somewhat better with the experimenters' plan than they had in *PART III*. So too there was considerably higher agreement in judgments on goodness and poorness of designs—this doubtless due, in part at least, to the impossibility of recording medium judgments. Average times for seeing out the part in the whole were as follows: for the six good designs 12.82″; six poor designs 12.48″; six designs with natural parts 3.30″; six designs with unnatural parts 21.99″; three good designs with natural parts 3.73″; three good designs with unnatural parts 21.91″; three poor designs with natural parts 2.87″; three poor designs with unnatural parts 22.08″. Thus the results obtained from this group support the conclusion drawn from the results with the other groups that a natural part of either a good or a poor design is seen out more easily than is an unnatural one. The results do not, however, show that parts, regardless of their nature of natural or unnatural, are seen out more easily in good designs than in poor designs; they do show it for unnatural but not for natural. As only a slight superiority on this point had been shown in the other experiments, it may be concluded from the additional data of *PART IV* that it is a factor of little importance.

Discussion and Conclusions

In our introductory statement we raised the question as to whether or not designs are seen as wholes made up of parts and whether a unified (good) design is more or less apt to be so seen than is a poor one. As our method called primarily for the seeing out of parts our results do not allow us to answer the first part of this question, although the fact that the parts of certain designs were seen as quickly as in one or two seconds does suggest that the seeing of parts in a whole is not an unaccustomed perception. As to the second part of the question, on the basis of our results we may say that, when the

naturalness of the part is taken into consideration, a good design has very little, if any, superiority to a poor design in the matter of seeing the part in the whole. What we called natural parts both of good and of poor designs were very easily seen, while what we called unnatural parts in both types of design were seen with much more difficulty. Naturalness of part, then, not goodness of design, is the more important factor in determining whether or not we will see parts in a whole. Coherence or unity in a design does not correlate positively with perceptual loss of separateness of parts, since designs judged coherent by a large majority of the Os had their parts very easily seen in the whole—providing these parts were natural. So, too, lack of coherence in a design does not correlate positively with loss of separateness of parts, since in these designs lacking coherence parts could also be very easily seen—provided they were natural.

To prove, as we think we have, that the perceptual coherence of a whole does not depend upon a loss of the sense of parts might seem a gratuitous endeavor—so decidedly from common observation one might expect this—were it not that a number of recent writers have contended that a primary law of perception is that of tendency toward 'monotonous form' or 'minimum variations.' Purdy[2] states that any factor which introduces variation interferes with coherence and that coherence is favored by homogeneity or monotony, that *e.g.*, a figure of solid color is more coherent than any other kind of figure. But coherence, the writer contends, is not merely a matter of homogeneity but just as much a matter of separate differing parts bound by means of certain principles into strongly coherent wholes. If degree of coherence depended upon degree of homogeneity then, *e.g.*, all designs in one color would be more unified than those in two or more colors. The telling use of color contrast or harmony in designs speaks against this assumption. If, then, high coherence can be found in wholes made up according to certain principles of different, separate parts, what are these principles? Our research affords only an indirect answer to this question. Designs which we constructed according to the æsthetic principles of tonality, balance, proportion, rhythm, harmony, nodality gave judgments of goodness (unity or coherence), while designs constructed without these principles gave judgments of poorness. It would seem then that these principles lend high coherence to a design even when it lacks homogeneity.

In our preliminary statement we also raised the question as to whether the parts of coherent (unified) designs possess certain characteristics and whether or not these same characteristics belong to the parts of designs not coherent. To answer this question completely in regard to the designs used in our experiment one would have (1) to determine just what are the parts in these designs; (2) to determine what are the characteristics of each of these parts. This we did not do. But since we *did* work with parts which we regarded as natural and as unnatural we can determine the characteristics which belonged to these 2 kinds of parts and also by determining which of these are common and which not, to the natural and unnatural parts of good and poor designs respectively, can indirectly throw light upon the characteristics of the parts of good and of poor designs.

[2] Purdy, D. M. The structure of the visual world. III. The tendency towards simplification of the visual field. *Psychol. Rev.*, 1936, 43, 59–82.

In the following discussion an attempt is made to characterize in a number of ways the parts of the twelve designs used in Parts III and IV. In analyzing these parts account has been taken of the fact, shown by a number of investigators,[3] that figures a little off regularity, symmetry, etc., will yet be seen as such—*e.g.*, although the parts in designs 3 and 8 are not physically exactly regular, yet they are, the writer believes, so seen perceptually. Choice of characteristics to be included in the discussion was made on the basis of those that have been emphasized in the literature as belonging to figures dominant to vision as well as those that the writer has found to be such. The characteristics studied, together with the definitions we gave them, were: (1) regularity (Reg), the enclosed space regularly bounded; (2) symmetry (Sym), figure divisible into two parts such that when figure is rotated on an axis one part may take the place of the other; (3) repetition (Rep), shape but not necessarily direction repeated; (4) good form (Form), all lines accepted as belonging together to form a stable whole; (5) complexity (Com), number of lines in the boundary of the part; (6) enclosure (Enc), closed shape; (7) isolation (Isl), determined by the number of lines in the part that continue beyond the part into the whole (where a line continues in one direction a value of 1 is assigned, where in two a value of 2, etc.; the highest degree of isolation therefore has a value of (0); (8) ratio (Rt), percentage of lines in the part to those in the whole; (9) familiarity (Fam), figure estimated to have been frequently seen; (10) proportionate division (Prop. Div.), part seen as a proportionate division of the whole; (11) tonality (Ton), part has characteristics important in setting the tonality of the whole.

Study of the designs shows that what we termed natural parts all have the characteristics of Reg, Sym,[4] Form, high degree of Isl (1.8 average as compared with 10.1 in unnatural), Fam, while none of the unnatural parts has these. The natural have a little less Com than the unnatural, the former averaging 5.8 lines with a range of 4–8, the latter 7.3 with a range of 5–9. The ratio of lines in the part to those in the whole seems not significantly different in the case of natural and unnatural parts, the former averaging 48.1 percent with a range of 23.5–66.6, the latter 54.5 percent with a range of 41.6–70. The natural parts of good designs have Rep, both physical and perceptual, while the unnatural parts of good designs, although they have physical Rep, seem not to have perceptual Rep. Unnatural parts of both good and poor designs lack Rep, both physical and perceptual, as do also the natural parts of poor designs.

Thus on the basis of the above analysis it can be said that, as far

[3] *Cf.* Koffka, K. *Principles of Gestalt psychology*, New York: Harcourt, Brace, 1935, 139 ff.

[4] The part of design 3, although on careful inspection lacking Sym, yet appears symmetrical.

as our experiment indicates, natural parts of good and of poor designs have several characteristics in common, and that these characteristics are lacking in unnatural parts of good and of poor designs. Also that the unnatural parts had slightly more complexity and a little higher ratio of lines of part to those of whole than did the natural.

Study also revealed that the natural parts of good designs have certain characteristics,—namely Rep, Ton, Prop. Div—which both the unnatural parts of good designs and the natural and unnatural parts of poor designs lack. It is significant that these characteristics are ones upon which certain of the structural principles of good design rest, namely rhythm, tonality, proportion. Without repetition of parts, rhythm in a design is impossible; and one of the easiest ways of accomplishing tonality in a whole is to have the same shape-feeling present in the part. In designs 1 and 5 it is quite evident that the shape characteristics of the part are those dominant in the tonality of the whole, while in design 11, although the shape of the part largely sets the tonality of the whole, it does so less exclusively than do the parts of 1 and of 5. Proportion when applied to a design means that the parts seem well proportioned to one another, and, as in our good designs with natural parts, this means that the parts cut up the whole space into proper divisions, or, in other words, fall at proportionate divisions. Had we possessed sufficient knowledge of harmony and nodality to have included them in our analysis we might have found that natural parts of good designs also have exclusive possession of these. However, we did not feel that just what shapes are harmonious and just what points in a spatial field are nodal had been sufficiently determined experimentally to warrant such inclusion. In constructing the good designs we did attempt to exemplify these two principles but often this is possible even without knowing upon just what such exemplification rests. And, as we have said, there has not yet been sufficient experimentation upon this point. Balance—the sixth of the structural principles of good design—refers, of course, to the balancing of parts and so could not be assigned to parts as such. We did find that the natural parts of good designs possessed symmetry, but so too did those of poor designs. But symmetry in the parts does not guarantee balance in the whole, since symmetrical parts may be placed in ways that do not effect balance, and often in design unsymmetrical parts are made to balance. However, if one looks at the natural parts of the good designs we used, one will see that these parts are instrumental in the balance of the whole, while both the natural parts of poor designs and the unnatural parts of good and of poor designs are not thus instrumental.

In the light of the analysis of the foregoing paragraphs it is indi-

cated that natural parts of good designs not only have characteristics in common with natural parts of poor designs but that they have other characteristics not thus common. Moreover, it is evident that these latter characteristics are those which assist in the structural composition of the design as a whole. If this is true, it raises the question whether or not what we called natural parts of poor designs were rightly so called. If so they are natural, not in the way of structural members, but in some way to be defined by means of the characteristics which we found common to all those parts which we called natural. That there was some ambiguity of meaning in the term natural as applied to parts of poor designs is suggested by the fact that in a number of the poor designs the Os did not agree in their judgments of naturalness of part; only after a number of substitutions of designs did we find three in regard to which the majority would agree, and in the final series there was disagreement among the Os as to the naturalness of part in two of the poor designs. Perhaps some of the Os felt, as does the writer after this research, that natural should mean structural in a sense in which the natural parts of our poor designs did not mean it. At least, the natural parts of good designs are natural in ways additional to those in which the natural parts of poor designs are. It may even be that parts of good designs possessing these latter factors but lacking those characteristics which we found present in the parts we used would yet be called natural. Certainly if they could be found, they would be structural. Whether natural *structural* parts must possess some or all of the characteristics we found belonging to the natural parts we used in our experiment is a question that should be put to experimentation. Only in this way can some very important questions in regard to design be answered.

It may seem in the light of the preceding discussion that our statement that "it is the naturalness of the part which determines whether or not we shall see it in a whole" should be changed to "it is the possession of certain characteristics which, etc." This may be true. But on the other hand we must remember that the Os' judgments of naturalness of part did correlate positively with their ease of seeing it as a part, and also that the more truly structural naturalness was present in the natural parts of good designs as we used them. As is evident, more research is needed to clear up the meaning and proper usage of the terms natural and structural in this connection and, if they are different, as seems probable, to determine their comparative roles in the perception of parts in a whole.

Brief Comparison with Some Related Researches

In recent years a rather large literature dealing with the problem of the principles governing the organization of a visual field has grown up.[5] Wertheimer, in a now classic article published in 1922, pointed out a number of these principles, such as nearness, similarity, enclosure, continuation of line, 'good form,' etc. A number of researches by Gottschaldt, Kopfermann, Gemelli, Galli, Gatti, Suzurki, Purdy, Foley, and others followed. Among these experimenters there is fairly large agreement in regard to at least certain of the principles, but there is a split among them as to whether the functioning of such principles is due to past 'experience' or, as the Gestalt psychologists contend, to 'natural' organization. Wertheimer spoke strongly against experience as the determining factor and Gottschaldt [6] performed an experiment which he thought verified this contention. He built up an 'experience' factor for certain figures by presenting them to his Os over five hundred times. These figures were then presented as parts of more complex figures and the fact that his Os did not 'see out' of the whole these part figures, despite the fact that they were thus loaded with the 'experience' factor, made him conclude that natural organization rather than experience determined the organization of a visual field. This experiment has met with wide discussion and much adverse criticism. It has been said, *e.g.*, that Gottschaldt did not take account of the fact that the complex figures may have had an even greater 'experience' loading than did the parts, even after they had been presented over five hundred times, and would thus resist being broken up. But Gottschaldt measured the time it took to 'see out' the parts in his various complex figures and the part figures were sometimes seen out of the more familiar complex figures in a shorter time than in the less familiar ones.[7]

In a more recent experiment Djang,[8] taking her point of departure from Gottschaldt's work, performed a learning experiment using dotted line figures, twelve simple, and twelve complex figures in which the simple figures were 'masked.' Contrary to Gottschaldt she concludes on the basis of her results that past experience has a decided effect upon visual perceptual organization. "The masked figures were seen as separate units in their respective complexes twenty

[5] For a good review of this literature: *Cf.* Foley, J. P. The effect of context upon perceptual differentiation. *Arch. Psychol.*, N. Y., 1935, 184, 5–67.

[6] Gottschaldt, K. Über den Einfluss der Erfahrung auf die Wahrnehmung von Figuren. I. Über den Einfluss gehäufter Einprägung von Figuren auf ihre Sichtbarkeit in umfassenden Konfigurationen. *Psychol. Forsch.*, 1926, 8, 261–317.

[7] *Cf.* Koffka, *op. cit.*, 155 ff.

[8] Djang, S. The role of past experience in the visual apprehension of masked forms. *J. exper. Psychol.*, 1937, 20, 29–59.

times more often by the experimental groups than by the control groups and seventy-seven times more often on the first trial. This great difference existed as an outcome of the former's having learned these figures in isolation." [9]

Although Djang's [8] results show that with her designs past experience with a simple design did have in many cases a large influence upon the subsequent perception of a complex design in which this small design was a part, still very significant differences in the degree of effect on the several designs are also shown. She gives in Table III, p. 42, the total number of successes made by all Os for each of the twelve designs used. This success varies from 13 to 91 percent. Since all of the small designs had been equally imprinted, in that they were presented until they could be correctly reproduced and thus in a very genuine way their 'experience' value had been equated, this great difference in the degree to which they were then seen in the complex figures in which they were masked must be accounted for by other than the 'experience' factor. Djang herself emphasizes this when she writes, "This variation in the role of past experience exists doubtlessly as a result of the relative ease of dissolving the different test-figures into the simple figures and the parts making up the camouflage." [10]

That experience played a considerable role in Djang's results cannot be denied and, considering the nature of the designs, was to be expected. These designs were conglomerate nonsense ones constructed as far as this writer can see with no regard for structural principles of design. They lacked unity and coherence in any legitimate use of these terms and in not a single case was the simple design a 'natural' (in the sense of structural) part of the complex in which it was masked. Moreover, what in our own experiment was shown to be important, the simple designs varied in the degree to which they had characteristics which would make them easily seen. Take, *e.g.*, the factor of isolation. In designs 3, 6, 10, 12—the four that Djang says are the easiest in which to achieve 'success' [*i.e.*, where the simple is seen in the complex in which it is masked]—we find that 3 and 12 have no boundary lines continued; 10 has one line continued but in such a way that in order to see it as a continuation it must become a curved line, a thing contrary to the nature of the straight linedness it has as a part of the small figure (a simple, almost symmetrical figure), and hence not easily occurring; 6 has two boundary lines continued but at points where continuation is clearly intercepted by the strong organization of the part figure. In figures 1, 9, 11—those most difficult of success according to Djang—in 1 and 9

[9] Djang, S. *Op. cit.*, p. 58.
[10] Djang, S. *Op. cit.*, p. 44.

there are two strong conspicuous continuations of boundary lines and in 11 clearly one and possibly two.

In order to see what results would be secured with Djang's figures by using a method more comparable to that of Gottschaldt the writer presented one at a time each of the simple, together with its complex, figures to a group of seventeen Os and timed them to see how long it would take for them to 'see out' the simple in the complex. It was thought that without *any* experience with the designs there might be differences in time of 'seeing out' that would in part at least account for Djang's differences in 'success.' No complete correlation was found although the following facts may be of some significance. The average time for 'seeing out' the part in those five designs in which Djang reports a 50 percent or better success was 5.79″ while for the remaining seven with below 50 percent success it was 7.72″. Four of the five with over 50 percent success were among the first five 'seen out' most quickly and their average time was 5.04″. On the other hand No. 6, in which Djang reports the greatest success, was the 9th in order of time to 'see out' and No. 9, which gave only 20 percent success, was 3rd in order of time to 'see out.' Although the above figures do not, as has been stated, show a complete correlation, still, considering the small number of Os, that four designs out of five with over 50 percent success were seen out most quickly does seem to indicate that something in these designs made them even without an 'experience loading' easily perceived in the complex designs of which they were part. That they were not more often so perceived by Djang's control group seems surprising. It is true that of the twenty-one times that a simple figure was noted in a complex by this group, nineteen of these occurred among the five having success over 50 percent. Since Djang's Os approached the designs with instructions to learn so as to be able to reproduce, it may be that 'natural sensory organization' had less chance to function.

It would be expected that in the learning of the whole design any part that had been previously seen and learned would be taken as the central unit around which the rest would be organized. To this writer one of the outstanding facts in Djang's experiment is the relatively small number of times this occurred in so many of the designs. That it did not is probably due to the fact that in these latter designs organizations occurred spontaneously according to natural sensory principles in such ways that the part units, previously learned, could not now function as such. Where no such organizations occurred, a figure previously learned would serve as the pivot for the new organization.

That Djang's Os went about their learning task in ways that would favor arbitrary perceptions is borne out by her report of their

methods. The most frequent method was to get the general contour and then complete the internal parts, but a considerable number of Os started in some arbitrary way such as to begin with the left portion or to proceed down from the top. Our Os, to be sure, were instructed to 'see out' the simple in the complex and thus their perceptions may not have been organized as they would had they been merely passively perceiving. On the other hand, the time that it took to organize in such a way that parts could be seen out seems a fair way of measuring the naturalness of such perceptions.

To the writer it seems that much more experimental research is needed before we can say definitely to what extent the organization of our visual fields depends upon past 'experience' and to what extent upon natural principles of sensory organization. In her own research it is to be noted that the parts seen out most quickly were supposedly more familiar than were those less quickly seen. But as our analysis shows the former possessed many factors other than familiarity which the latter lacked, why attribute to one factor—familiarity—the dominant role? One could hardly say that we see out the regular, the symmetrical, etc., because these are the most familiar. That they are is surely contrary to our experience. Moreover, as has often been pointed out by the Gestalt psychologists, if there are natural principles of organization one would expect man-made articles to be constructed according to these and hence the naturally organized would come to be frequent. For this reason it would be difficult to find material for experimental research that would be able to test the relative weights of familiarity and natural organization. But such a study should be made. A finding in regard to one of our designs—number 5—is perhaps relevent here. In *Presentation 1, PART I*, the part drawn by a large majority of the Os was the double-triangle. Why was this drawn instead of a parallelogram (the part which, it may be added, had been chosen by the experimenters as the natural part)? Surely the latter is at least as familiar as the double-triangle. And the same thing may be said for the central diamond.

Two other researches should be mentioned in this connection. Hollingworth [11] in an experiment on Verbal Configuration found that words could be built from letters presented as a meaningful word as quickly as from the same letters presented as a meaningless word. This he interprets as speaking against the influence of the Gestalt. Since according to his assumption a meaningful word is as good a Gestalt as can be found and since in his experiment he found it ineffective he concludes that the influence of the whole upon its parts

[11] Hollingworth, H. L. The conditions of verbal configuration. *J. exper. Psychol.*, 1935, **18**, 299–306.

is negligible. And he goes on to criticize the Gestalt principles of closure and completion, finally stating, "We are disposed to carry the conclusion over to all the other cases in which closure, completion, supplementation, have been introduced as explanatory terms. Words, we have concluded, are not Gestalten. But they come as near being this as does anything else." [12] To the writer this seems a very easy way to dismiss, without presented critical evaluation, a large number of very careful experimental researches. It may be, as Hollingworth says in discussing reports on the filling in of the blind spot, that "people report more than they see," but if such reports are not to be trusted much that in the past has been accepted by psychologists, of other than the Gestalt school, will have to be thrown out. What Hollingworth's study suggests to the writer is that a meaningful word is not a very strong Gestalt. But this does not prove that such strong Gestalten do not exist—such Gestalten, *e.g.*, as the good designs used in the present experiment. It may well be that wholes organized according to 'natural principles of organization' are much stronger and have more influence upon their parts than do more arbitrarily constructed ones, such as seen words. If these, by the way, had any Gestalt quality it could only come from familiarity, not from principles of natural sensory organization.

That meaning alone will not give strong visual coherence to a configuration is also suggested by a study by Foley.[13] In the first part of his experiment Foley had his Os differentiate into parts, according to a scheme given them, 4 sheets of designs, two consisting each of twelve pictures of meaningful objects, and two of twelve meaningless or nonsense designs. Differentiation in the case of the meaningful designs was found more difficult. This seemed to suggest that meaning impeded differentiation. But study of the Os' introspections suggested that another factor—'compactness'—was at work. Compactness denotes "the degree to which adjacent masses of black or white occurred in the total design. Thus a relatively 'compact design' would be one in which the black areas clustered together in a more or less solid mass with minimal contour." [14] Study of the designs showed that "the so-called meaningful designs are reliably more 'compact' than the corresponding meaningless designs." [15] Three sheets of designs (one of compact and meaningful, one of compact and meaningless, one of non-compact and meaningless) were then differentiated by a new group of Os. Differentiation in the case of the 'compact' designs, both meaningful and meaning-

[12] Hollingworth, H. L. *Op. cit.*, p. 306.
[13] *Loc. cit.*
[14] Foley, J. P. *Op. cit.*, p. 36.
[15] *Op. cit.*, p. 37.

less, was harder than in the case of the non-compact, meaningless designs. It would seem then that 'compactness,' not meaning, was the factor that impeded differentiation. To verify this a third experiment was performed in which by drawing a number of horizontal lines across each of the designs, all three sheets were accepted as being equated as to 'compactness.' The results of the Os showed 'no significant differences in mean score between any of the three forms.' [16]

Having accepted 'compactness' as the causal factor in his results, Foley then writes—"This contextual factor of 'compactness,' however, was not a function of the entire black figure, as would be expected in accordance with orthodox Gestalt theory. The difficulty of differentiating the parts cannot be attributed to the influence of the total figure. . . . The total figure is not changed in 'effect' by the introduction of the fine horizontal lines. . . . If the total 'configuration' were the factor impeding differentiation, these fine horizontal lines should not make any appreciable difference." [17] Foley goes on to say that 'compactness' might be interpreted in several ways, *e.g.*, "according to such an orthodox Gestalt viewpoint, the important role of 'compactness' . . . might be described in terms of 'proximity,' 'similarity,' 'continuity,' '*pragnanz*,' and the like . . . but it has been experimentally shown in the present investigation that such contextual factors as are operative do not reside in the total pattern." [18] He goes on to suggest that 'compactness' impeded differentiation because "long experience has familiarized us with innumerable small and 'compact' or homogeneous objects, the parts of which have been found to belong together through empirical interaction." [18]

To the writer it seems that Foley's finding in regard to 'compactness,' far from disproving the role of principles of organization pointed out by the Gestalt psychologists, is a verification of certain of them. To her knowledge no Gestalt psychologist has stated that a visual field organizes according to characteristics belonging to the whole field. Rather, certain principles function to make the field organize in specific ways—*e.g.*, similar dots in a visual field will group together to form certain separate units—but no one supposes that 'similarity' belongs to the whole field. Because of the law of 'good continuation of line' certain lines, rather than others in the field, will form a unit. But good continuation does not characterize the whole field. Both similarity and continuity of line seem to the writer to underlie 'compactness.'

[16] *Op. cit.*, p. 58.
[17] *Op. cit.*, p. 61.
[18] *Op. cit.*, p. 62.

It is to be regretted that Foley gives no data in regard to the individual designs in his various sheets, since a study of these might reveal factors other than that of 'compactness' functioning to effect the results. Foley states (p. 43) that the designs of Form A lacked symmetry and since Form C was constructed from figures of Form A, supposedly the same holds here. To the writer's perception some of the designs of these forms were symmetrical, some not. To check her perception she presented designs 4, 5, 8, 11, 14, 15 of Form A (Exper. II, p. 40) and 3, 5, 6, 11, 13, 15 of Form C (Exper. II, p. 42) to a class of twenty-one Os and found a 100 percent agreement on judgments of symmetry for 5, 11, 15 in Form A and 5, 11, 15 in Form C and a 100 percent agreement on judgments of non-symmetrical for 4, 8, 14 of Form A and 3, 6, 13 for Form C. It may well be that the factor of symmetry as well as that of 'compactness' was at work in effecting Foley's results. Without a prior equating of the several structural principles of the separate designs it does not seem justifiable to throw them together into groups and then single out one factor as the causal one in effecting the results. A study of the various forms used does reveal that on the whole the designs of some of these are more 'compact' than are others but since this is not true of all of the individual designs, to attribute the difficulty of differentiation to this one factor seems unjustifiable. Also in the last part of Experiment 2 not only is 'compactness' done away with but doubtless other structural principles as well. This is why the writer cannot agree with Foley's statement that "The total figure is not changed in 'effect' by the introduction of the fine horizontal lines." She finds the 'effect' distinctly changed.

General Conclusions

1. Designs constructed with regard for the æsthetic principles of balance, proportion, rhythm, tonality, harmony, nodality, were judged good (unified or coherent). Designs lacking these characteristics were judged poor.

2. Degree of coherence in a design did not correlate positively with perceptual loss of separateness of parts.

3. Natural parts of both good and poor designs were readily 'seen out' of the whole, while unnatural parts were not.

4. Natural parts of both good and poor designs were found to have the characteristics of regularity, symmetry, good form, high degree of isolation, familiarity, while unnatural parts lacked these.

5. Complexity of the part and proportionate number of lines in the part and the whole seem not to be differentiating characteristics of natural and unnatural parts.

6. Natural parts of good designs also had the characteristics of repetition, tonality of the whole, proportionate division, while natural parts of poor designs lacked these. Unnatural parts of both good and poor designs also lacked these characteristics.

7. Natural parts of good designs were instrumental in effecting the balance of the designs, while the unnatural were not. Neither natural nor unnatural parts effected balance in the poor designs.

8. A distinction should be made between 'natural' and 'structural' parts of a design. The latter are probably those necessarily involving æsthetic principles while the former need not involve them.

9. Data from *Presentation 1, PART I*, suggest that natural parts are dominant to vision.

10. In the light of her data, consideration of other comparable researches makes the writer conclude that principles of natural sensory organization as well as—probably even more than—experience or meaning factors must be considered as determinants of our visual perceptions. Much more research is needed to determine their relative weights.

(Manuscript received August 10, 1940)

EXPERIMENTAL STUDY OF THE INFLUENCE OF VISION ON SOUND LOCALIZATION [1]

BY GARTH J. THOMAS

University of Kansas

Introduction

Studies of vision without retinal inversion such as Stratton's (7) and Ewert's (1), and also studies of binaural transposition such as P. T. Young's (11) and that of Willey, Inglis, and Pearce (10) have shown that often the visual perception of an object will affect the phenomenal auditory localization of the sound it makes. Willey, Inglis, and Pearce (10) say, "When the observer can see the source of sound, it was found in the experiments of Stratton, Young, and Ewert, and in the present experiment, that the sound sometimes seemed to be coming from the position of the seen object" (p.123).

When the problem of how perceptions in two modalities can influence one another is considered, the most prevalent answer, and the one psychologists have been most wont to give as the main factor in perceptual organization, is that the person must *know* that this particular visual object also makes this particular sound. Such an answer assumes that some 'experience' or 'learning' factor, unique to the individual, must be involved in causing the two experiences, though arising in separate modalities, to belong together. Undoubtedly past experiences of the individual lead to important factors which enter in his processes of perceptual organization. It seems equally plausible that these 'past experience' factors could be considered as essential in some intersensory phenomena as mentioned above.

However, Hornbostel (6) has proposed a dynamic concept of the unity of the senses. From such a concept, if true, it would be expected that intersensory relations can arise directly from the dynamic organizational processes in perceiving, and not mediately as a phenomenon based on 'past experience.' An intersensory phenomenon of this nature *not* arising from the experiential factors of the individual is derived from the dynamic perceptual organization of the organism on a phylogenetic level. For a history of the experimental work on intersensory relations see Hartmann (3), (4), and (5).[2]

[1] The author is indebted to Dr. Fletcher McCord for suggesting the problem.

[2] For an even more complete history of the problem see Ryan, T. A., Interrelations of the sensory systems in perception, *Psych. Bull.*, 1940, 37, 659–698. Ryan's article appeared after the present paper had gone to press. The present study gives confirmatory evidence for some of the findings of Klemm reviewed by Ryan (pp. 673–675).

The purpose of this study was to ascertain whether visual perceptions can influence phenomenal sound sources on a level of perceptual organization in which past experience factors are reduced to a minimum, *i.e.*, the attempt was made to discover whether this intersensory relation between vision and sound localization can be considered as a dynamic inter-relation between the sense fields. This was done by using as stimulus, configurations in the two modalities concerned, stimuli which were meaningless to the subjects in the sense that they were not familiar objects or patterns. The stimulus configurations used were those of rhythmic patterns.

The study had a secondary purpose besides the primary one described above. That purpose was to demonstrate an experimental procedure which may be of value in further studies on the problem of intersensory relations.

Description of the Apparatus

In general the set-up consisted of a curved screen on which were projected two spots of light at a fixed distance apart. Behind the screen was a movable sound source. The O saw the two spots of light on the screen, heard the sound, and judged as to whether the sound was closer to the left light, the right light, or the middle region.

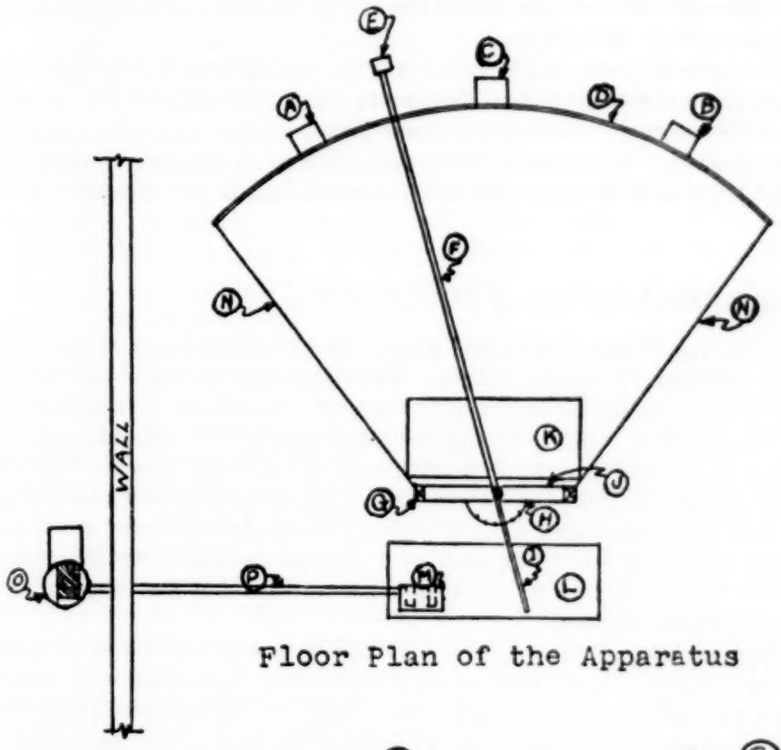
A more detailed consideration of the set-up is as follows: The screen was curved making an 85° arc of a circle four meters in diameter. The O sat in front of it (on the concave side) at a radius distance of two meters from its center. From one side of the screen to the other, *i.e.*, the chord of the arc, was approximately 2.58 meters. In its vertical dimension the screen was one meter, and it was so raised from the floor that the middle of its vertical dimension was approximately at eye level of a person sitting in a chair. The screen proper consisted of white sheeting tacked on a wooden frame. (See (*D*) in diagram, page 165.) Behind the screen, *i.e.*, on the opposite side from the O, were two light boxes (*A* and *B* on diagram, page 165). They consisted merely of light-tight boxes with a round hole of about 13.5 cm in diameter in their front ends. The holes were covered with frosted glass. The light was furnished by a forty watt frosted light bulb in each box. The fronts of the boxes were placed snugly against the screen so that from the front side of the screen two sharply defined, round spots of light with the same diameter as the holes in the boxes were visible. The lights were at eye level or at the center of the vertical dimension of the screen; however, horizontally they were somewhat more than 10° from each side of the center.

The sound source consisted of an ear-phone made to sound by a 60 cycle, eight volt current. A rheostat further cut down the current until the phone made a low buzzing sound. As the sound source was to be heard from different places behind the screen, some way of silently moving it about was needed. This was accomplished by having it suspended from a long arm or boom which turned on a pivot at two meters, or radius distance, from the screen (*E* and *F* on diagram, page 165). This pivot was directly above the O's head. The O sat between two uprights, and cross-pieces at a level higher than the O's head held the pivot or axis for the boom. The axle was a piece of one inch steel shafting. From this swung the arm, extending over the top of the screen. Suspended from it behind the screen at the same level as the lights was the sound source. By moving the arm back and forth, the sound source could be moved silently to any position between the two lights.

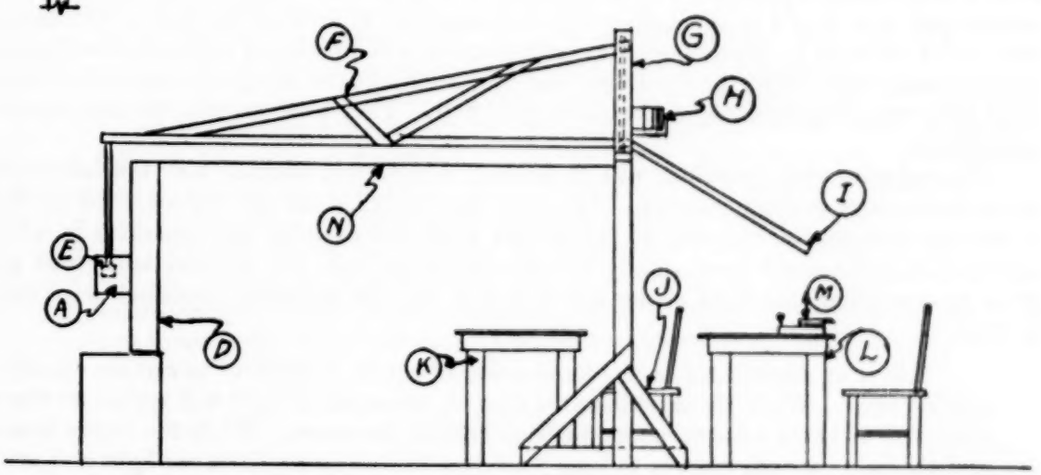
Directly in front of the O was a table to the edge of which was clamped a modified Hering head clamp. The clamp did not prohibit all head movements, but at least kept the O's head at a constant distance from the screen and also made it possible for him to hold his head relatively still. He could rest his elbows on the table and thus be comparatively comfortable during the experimental session. Directly in front of the O, in the center of the screen, equidistant from either light and from the top and bottom of the screen was a small ruby light which was the

fixation point for the O. It was made by having a small light-tight box with a light in it and with a tiny hole in the front covered with red cellophane. This was set snugly against and behind the screen.

From the two ends of the screen to the two uprights, on either side of where the O sat, were stretched two wires. Hanging over these wires and reaching from the top edge of the screen to the frame holding the boom was a curtain of black stage cloth. This formed a ceiling for the O above his head, but below the boom. Also black curtains were on either side. Thus the O found himself in a small wedge shaped room with himself at the point, the screen making up the



Floor Plan of the Apparatus



Side View of the Apparatus

DIAGRAM I. (*A*) Light box. (*B*) Light box. (*C*) Light box for fixation point. (*D*) Screen. (*E*) Buzzer. (*F*) Boom from which the buzzer is suspended. (*G*) Uprights forming framework for support of the boom. (*H*) Dial indicating boom position. (*I*) Stick by which the experimenter moved the boom. (*J*) Position of observer. (*K*) Table in front of observer. (*L*) Table for the experimenter. (*M*) Switches for controlling the lights and buzzer. (*N*) Wires from which cloth curtains were suspended. (*O*) Turntable acting as a current interrupter. (*P*) Cable of wires from the turntable to the switches.

periphery, and the ceiling and either side made up of black stage cloth curtains. Between the bottom of the screen and the floor was also suspended a stage cloth curtain.

The E sat at a table directly behind the place at which the O's chair was situated. (See *L* in diagram, page 165.) Extending backwards from the boom was a stick which projected out over the E's table by which he was able to move the boom back and forth. (See *I* in diagram, page 165.) On the lower one of the cross pieces between the two uprights was a small semicircle or arc whose center was the boom's axis. (See *H* in diagram, page 165.) On it were indicated the positions from which the buzzer was to be sounded, and a pointer from the boom indicated the boom's position. A small light illuminated this dial so the E could see it when the room was darkened. On the E's table were two switches which controlled the lights and sound source. The apparatus was in the experimental dark room.

As the stimulus configuration to be used was rhythm, a way of making that rhythmic pattern was needed. In the room adjoining the experimental dark room a turntable was set up. It was turned by a constant speed electric motor. A copper plate was placed on the turntable. A series of contact brushes pressed against it. By covering the copper plate with papers appropriately slotted, electric current to the lights and buzzer could be interrupted in any repeating rhythmic pattern chosen.

Experimental Procedure

There were four experimental conditions used in this study. In Condition I, one of the lights was on constantly during the presentation of the stimuli. The other light flickered off and on in the same rhythm as that of the intermittently sounding buzzer. Condition II consisted of having the sound and one of the lights on steadily during the presentation of the stimuli, and the other light flickering off and on in the same rhythmic pattern as it had in Condition I. In Condition III *both* lights flickered off and on, but *one* of them was always in rhythm with the intermittently sounding buzzer, and the other light flashed with an altogether different rhythmic pattern. The fourth and last condition was introduced as a control condition for the whole experiment. In this case, both lights flickered together in the same rhythmic pattern, but the buzzer sounded in a different rhythmic pattern.

The Os in this experiment were ten students, five boys and five girls, from the author's beginning psychology laboratory classes. They were selected on the basis of available free time from a larger group of students who volunteered to be subjects. They were all college sophomores taking their first course in psychology and consequently, at least at the start of the experiment, were quite naive as to psychological experimentation and problems; however, the experimental sessions lasted from about the third week until the middle of the term, so by the end of the experiment they were presumptively not so naive as at first. They were not told the true nature of the experiment.

The experimental procedure was as follows: The O was ushered into the laboratory and seated in his chair in the apparatus. The room was dimly lighted and the source of the light was within the little curtained room, so the subject could not examine the apparatus in detail. In approaching it he could have seen it in dim outline perhaps, but no close study was possible. After he was seated the head clamp was adjusted and the following directions were then read to him:

> This is an experiment in the localization of sound in relation to certain visually given space frames. When the experiment is started, two spots of light will appear on the screen. Also, you will hear a buzzer sounding from behind the screen. While the buzzer is sounding and the lights are on, you are to keep your eyes fixated on the little red light and your head as still as possible, *i.e.*, the clamp is not to prohibit you from making head movements but to give you a substantial rest so you can hold your head still. When the room lights go out (indicated by pointing out what was meant by 'room lights') and the screen lights come on, and the buzzer begins to sound, you are to judge whether the sound of the buzzer appears to be coming from nearer the left light, from nearer the right light, or from nearer the middle. When you make your judgments you will say 'right,' 'left,' or 'middle,' depending on where you experience the sound as being. When the stimuli are presented, wait until you get an impression of the sound's location, even though vague, and then make your judgment on the basis of that impression. *You* are not being tested, but rather you are acting as an

observer and your task is to report where *you* experience the sound of the buzzer in terms of being closer to one light or the other, or nearer the middle.

Try to keep your head as still as possible, but especially be sure to keep your eyes fixated on the red light while the stimuli are being presented. Between presentations you may move your eyes to rest them, but when the room lights go out, that is your signal to fixate the little light. If you should accidentally move your eyes just report it to me, and a new presentation will be given.

After reading the instructions to the O he was asked if he had any questions concerning his task, and if so they were answered, either by re-reading passages of the formal instructions, or by further explanation. The idea of the eye fixation was usually emphasized further.

In all of the conditions the buzzer was sounded from five different places. These stimulus positions are referred to as 1, 2, 3, 4, and 5 in the data. Position 1 was next to the left light or exactly 10° from the center of the screen. Position 2 was 5° to the left of center. Position 3 was exactly in the center, directly behind the fixation point. Position 4 was 5° to the right of center, and position 5 was 10° to the right of center, or next to the right light.

In all the conditions except number IV the side on which the 'in-rhythm' light (*i.e.*, the light flickering in the same rhythmic pattern as the buzzer) could be changed, thus controlling any effect position might have. In Condition I, for example, one light was on constantly and the other light flickered in the same rhythm as the intermittently sounding buzzer. The position of the flickering light, either to the right or to the left, was determined by merely throwing a switch. The particular position from which the buzzer was sounded in any one trial and which side the 'in-rhythm' light was on were determined by a prepared chance order schedule. Thus in any one condition each O made 200 recorded judgments, 40 with the sound at each of the positions, and of these 40 judgments, 20 were made with the 'in-rhythm' light on the left and 20 with it on the right—in a prearranged scrambled order, however.

The O was given 10 practice or unrecorded trials each time he went through a series, and each of the ten Os had all four of the conditions. The order in which the different Os went through the four different conditions was staggered as much as possible to control any effect the order of the conditions might have. Three of the Os had the order of the conditions as follows: I, II, III, IV. Two of the Os were presented the four conditions in the order III, IV, I, II. Three more of the Os had the order of presentation of IV, III, II, I; and the last two Os were presented the conditions in the order II, I, IV, and III. Thus I followed II the same number of times that II followed I, and III followed IV the same number of times that IV followed III. The pair III and IV were first with half of the Os and I and II were first with the other half of the Os. III and IV were considered sufficiently different from I and II to warrant pairing them that way. With only ten Os and four conditions that represents about the maximum staggering.

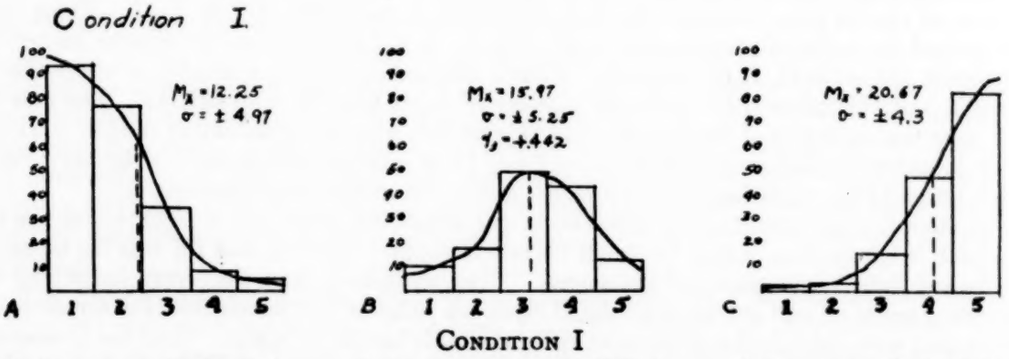
The stimuli were presented as follows: At the beginning there was a light under the O's table which brightly illuminated his small curtained room. Turning this off was his signal to fixate the red light in the center of the screen. Then an instant later the screen lights and the sound were turned on and left on until the O made his judgment at which time the screen lights and the sound were turned off and the light under the table turned on again. Then the E recorded the judgment and reset the position of the buzzer by moving the boom. The switch controlling the side on which the 'in-rhythm' light appeared was set and then the second presentation was started. No systematic introspections were requested, but at the end of each experimental series the Os were asked to write out their answers to a number of questions.

Results

In the treatment of the results of this study the combined data of all ten subjects are used. The results are represented graphically on pages 168, 169, and 170. Each set presents the distributions of judgments of one of the experimental conditions. Set I presents the results for Condition I. It will be remembered that Condition I consisted of having one of the screen lights flickering in rhythm with the intermittently sounding buzzer. The other light was on steadily.

Graphs *A*, *B*, and *C* represent respectively the distributions of the judgments called 'left,' 'middle,' and 'right' while the 'in-rhythm' light was on the *left* side of the screen. Graphs *D*, *E*, and *F* present correspondingly the 'left,' 'middle,' and 'right' judgments when the 'in-rhythm' light was on the *right* side. The vertical axis of the graphs represents the percentage of judgments, and the numbers on the horizontal axis represent the stimulus positions from which the buzzer was sounded.

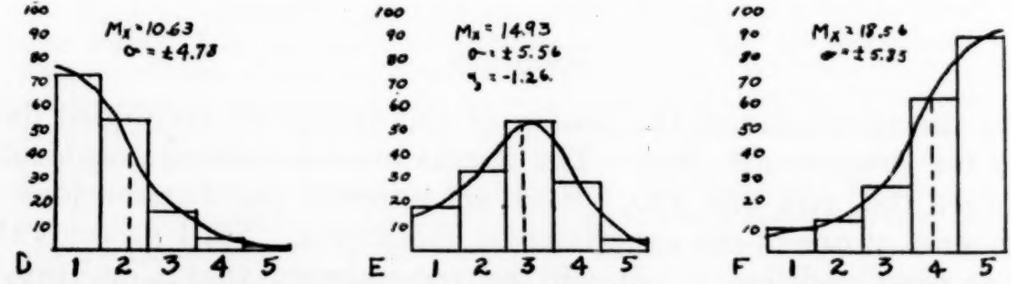
Set II shows graphs representing the results under Condition II in which the buzzer sounded steadily and one light flickered off and on, and the other one was on constantly during the interval of stimuli presentation. The individual, lettered graphs represent corresponding situations as in Set I. Set III presents the results under Condition III in which both lights flickered in different rhythmic patterns and the buzzer sounded intermittently in the same rhythmic pattern as one of the lights. The individual graphs represent corresponding situations as the above two. Set IV presents the results of Condition IV in which both lights flashed together in the same rhythmic pattern while at the same time the buzzer sounded intermittently in a different pattern. Graph *A* presents the distribution of 'left' judgments, *B* of 'middle' judgments, and *C* of 'right' judgments.



CONDITION I

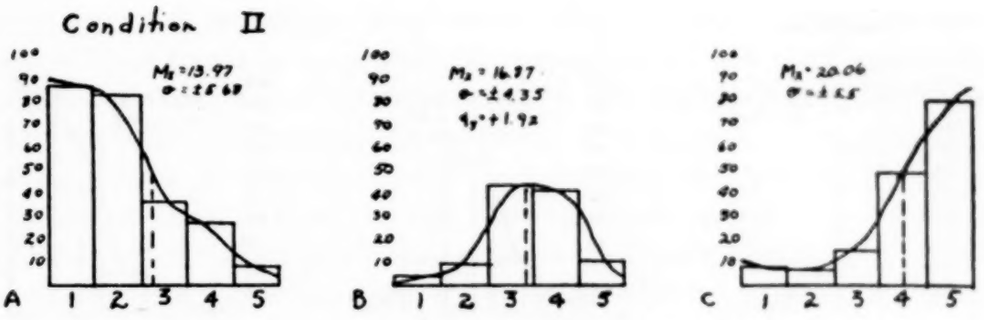
Judgments called 'left.' Judgments called 'middle.' Judgments called 'right.'

The above graphs present the distributions of judgments when the *left* light flickered and the *right* light was maintained steadily. The buzzer sounded in rhythm with the *left* light.



Judgments called 'left.' Judgments called 'middle.' Judgments called 'right.'

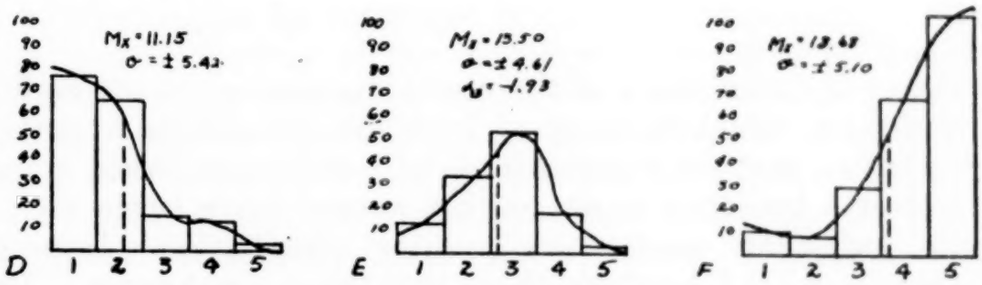
The above graphs present the distributions of judgments when the *right* light flickered and the *left* light was maintained steadily. The buzzer sounded in rhythm with the *right* light.



CONDITION II

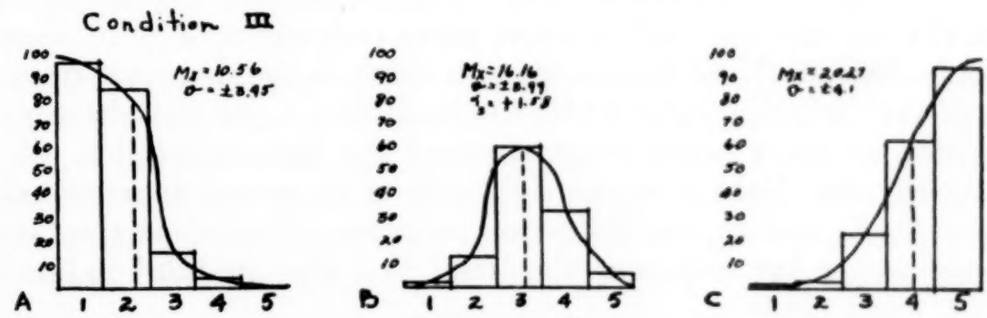
Judgments called 'left.' Judgments called 'middle.' Judgments called 'right.'

The above graphs present the distributions of judgments when the *left* light flickered and the *right* light was maintained steadily. The buzzer sounded steadily.



Judgments called 'left.' Judgments called 'middle.' Judgments called 'right.'

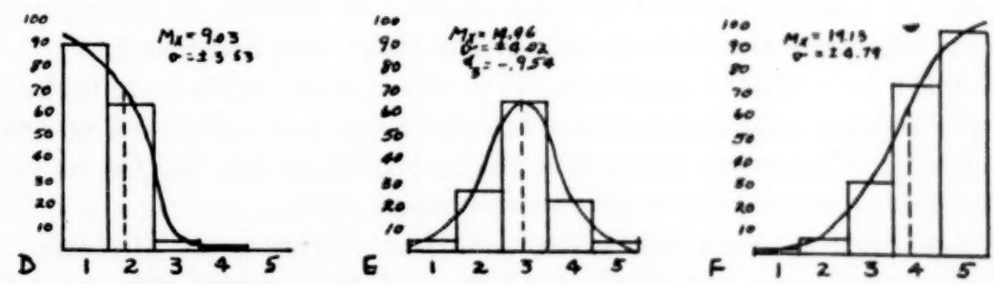
The above graphs present the distributions of judgments when the *right* light flickered and the *left* light was maintained steadily. The buzzer sounded steadily.



CONDITION III

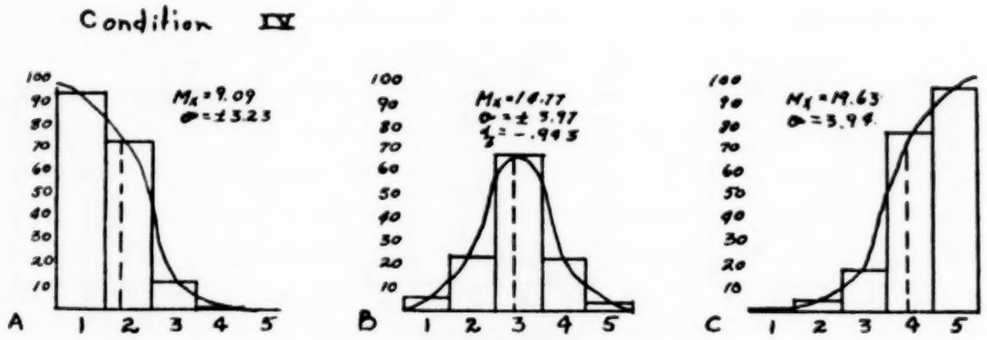
Judgments called 'left.' Judgments called 'middle.' Judgments called 'right.'

The above graphs present the distributions of judgments when both lights flickered but in different rhythmic patterns from each other. The buzzer sounded in rhythm with the *left* light.



Judgments called 'left.' Judgments called 'middle.' Judgments called 'right.'

These graphs present the distributions of judgments under identical conditions as those above except, in this case, the buzzer sounded in rhythm with the *right* light.



CONDITION IV

Judgments called 'left.' Judgments called 'middle.' Judgments called 'right.'

The above graphs present the distributions of judgments when both lights flickered in an identical rhythmic pattern. The buzzer sounded intermittently but in a different rhythm from the lights.

If there is no influence of the visual stimulation on the auditory localization, it would be expected from the phi-gamma hypothesis, *ceteris paribus*, that the judgments of 'left' and 'right' when graphed would show a tendency to approach a normal ogive curve with the 'middle' judgments approaching a normal distribution. Inspection of the graphs of Set I, however, shows that this is not the case. There appears a definite tendency for the Os to skew their judgments of the sound's location toward the side on which the light is flickering in the same rhythm with the buzzer. That is, when the flickering light is on the left the Os tend to make more judgments of 'left' than of 'right.' 'Middle' judgments tend to skew in the direction opposite that of the flickering light which also indicates a general phenomenal distortion of the sound's source toward the flickering light. When the 'in-rhythm' light is on the right side of the screen represented by graphs *D*, *E*, and *F*, the distortion reverses. This time the 'right' judgments are favored over the 'left' and the 'middle' judgments skew toward the left.

In Set II, which shows the results of Condition II in which the sound was on steadily and one light flickered and the other light remained on constantly, the same tendency is present as appeared in Set I. The Os tended to experience the source of the sound as distorted toward the side on which the light was flickering.

Set III represents Condition III in which *both* lights were flickering and the buzzer sounded intermittently in the same rhythm as one of the lights. This time there is a slight tendency for the Os to judge the sound source as closer to the *in-rhythm* light.

Set IV represents the results when both lights were flickering together in the same rhythmic pattern and the buzzer was sounding in a different rhythmic pattern. In this case, as the effect of the lights would be neutral, the distributions of the judgments should be

symmetrical. They come very near being this, but close inspection will show that there is a very slight tendency for the Os to skew their judgments of the sound's source toward the right side of the screen. At first it was believed that this was a constant error caused by standing waves in the room, but on plotting the individual records it was found that five of the Os skewed their judgments to the left and five of them skewed them to the right. As the combined data show, evidently those skewing their judgments to the right did so to a greater extent than those who skewed their judgments to the left. With a greater number of Os such a distortion should disappear.

The consistency with which the tendency for the sound source to be heard closer to the flickering or in-rhythm light, as the case may be, is evidence for some sort of constant relationship. Nevertheless further analysis of the data was undertaken to determine whether differences caused by the lights were statistically reliable. Means and standard deviations were determined for these distributions by the Spearman Arithmetic Method for the Method of Constant Stimuli. These were calculated by considering the numbers representing the stimulus positions to be the midpoints of class intervals. Thus the mean of graph *A*, Condition I, representing 'left' judgments with the 'in-rhythm' light on the left came out to be 12.25°. The mean of graph *D* of Condition I which represented the 'left' judgments with the in-rhythm light on the right came out to be 10.63°. The values for each distribution of judgments are found on the corresponding graphs. Critical ratios of the differences between various means of these distributions were calculated. In Set I the first critical ratio was between the means of graphs *A* and *C* to discover whether there is a reliably greater number of 'left' judgments when the flickering light is on the left than there are 'right' judgments. In order to compare them, one of the distributions was reversed so as to compare corresponding stimulus positions. When this reversal is made of graph *A* presenting the 'left' judgments and the mean is calculated, it was found to be 17.75°. The mean of graph *C*, presenting 'right' judgments, was 20.68°. Thus, there is a 2.93° difference in favor of the in-rhythm side. The critical ratio of this difference was 4.5 which is well above that considered to be statistically reliable.

When the critical ratio of the difference between the mean of the distributions of graph *D* presenting 'left' judgments with the in-rhythm light on the right side of the screen, and graph *F* which presents the 'right' judgments with the in-rhythm light on the same side, it was found to be only .98. This is definitely not considered reliable. None the less there is an increment of right side judgments over left side judgments of corresponding stimulus positions.

In comparing the two means of the two distributions of the 'middle' judgments, distributions *B* (in-rhythm light on left) and *E* (in-rhythm light on right), a critical ratio of 2.3 was found. A coefficient of skewness was worked out for the middle judgment distributions. The coefficient alpha$_3$ was used, *i.e.*, the third moment about the mean divided by the standard deviation cubed. For graph *B* the coefficient of skewness was +.44 and for graph *E* it was −1.26.

Two other comparisons were made and their critical ratios computed. The mean of graph *A* presenting 'left' judgments with the in-rhythm light on the left was compared with that of graph *D* which presented 'left' judgments with the in-rhythm light on the right. This comparison was to determine whether the side on which the in-rhythm light appeared made any difference in the 'left' side judgments. The critical ratio of the difference was 2.1. With graphs *C* and *F* presenting 'right' side judgments with the in-rhythm light on the left and right respectively, the critical ratio was 2.7.

Similar comparisons were made on the remaining sets of graphs. The various critical ratios can be read from Table I. They

TABLE I

CRITICAL RATIOS OF DIFFERENCES BETWEEN THE MEANS OF THE VARIOUS DISTRIBUTIONS

Graphs Compared	Set I	Set II	Set III	Set IV
A and *C*	4.5	4.7	1.5	2.5
D and *F*	1.0	.19	3.0	—
B and *E*	2.3	8.1	4.6	—
A and *D*	2.1	3.4	3.0	—
C and *F*	2.7	1.6	1.8	—

show a wide variation and about half of the differences are large enough to be considered definitely reliable. In all of the sets of data there is found a tendency, even though often slight, for the Os to judge the location of the sound closer to a flashing light, whether the sound is intermittent or not. When both lights are flashing in different ryhthmic patterns and the buzzer is sounding in an identical rhythmic pattern as one of the lights, the Os tend to judge the sound as closer to the in-rhythm light. Even though all the critical ratios do not indicate reliable differences, the *consistency* with which the skewness favors the flickering or in-rhythm light is good evidence that there is some influence of the light on the sound's localization and that the results do not indicate merely chance variations.

INTERPRETATION AND CONCLUSIONS

In order to interpret these results as indicating a dynamic interrelation of the sense fields, another possibility must be examined.

It will have to be shown that these results are not the result of an experience or learning factor.

In the first place it seems rather doubtful that a flickering light and an intermittently sounding buzzer could have any very definite experientially established meaning. There is no apparent reason why the subject should say to himself, "This light and this sound have the same rhythm, therefore they must belong together." This type of stimulus is even more 'meaningless' than geometric forms used in studies determining the *gestalt* factors in perceptual organization. A rhythmic pattern which is an unusual and temporal *gestalt* is hardly 'like' anything in the O's experience.

Further evidence that experientially derived meaning factors did not make the Os consciously skew their judgments of the sound's location toward the in-rhythm light can be seen in the answers to questions the Os wrote at the end of each experimental session. After each session they were asked whether they thought the lights influenced their judgments. After the presentation of Condition I six of the Os definitely said that the lights had no influence on their judgments. The other four reported that the lights *may* have influenced their judgments, especially when they were not sure. Examination of the individual results, however, showed no noticeable differences between those who reported the lights as influencing them and those who did not. After Condition II, only two of the Os reported any influence of the lights on their judgments. Two other Os showed skewness of their judgments toward the steady light rather than toward the flickering one. All the Os reported more uncertainty in this condition. Condition III in a sense represents a critical situation for this experiment because with *both* lights flickering the factor of the attention value of the flickering light as causing the results is eliminated. After being presented this condition again two of the Os reported that the lights may have influenced them in their judgments. In Condition IV no O reported being influenced by the lights.

After Conditions I and III the Os were asked if they noticed the buzzer and one of the lights were in the same rhythmic pattern. After Condition I, only one reported noticing the light and sound had the same rhythmic pattern. After Condition III two Os reported noting the identity of rhythmic patterns; however, neither of them was sure whether it was in rhythm all of the time or only part of the time. They had not clearly noted the relationship. When asked concerning how they interpreted their task, all of them reported that they thought the lights were to distract them; and they consciously took the task-set of disregarding the lights. This was true in all of the conditions. Thus all of the Os report a negative attitude from

the experimenter's point of view, and yet they did show a consistent distortion of their judgments of the sound's source toward the flickering or the in-rhythm light. The probability that conscious associations or experientially established meanings which might have been projected into the situation caused the results is very slight.

Another factor which might be used to explain the results is that of eye-movement. Goldstein and Rosenthal-Veit (**2**) have shown definitely that eye movements do influence sound localization. However, in this study the attempt was made to eliminate this factor as much as possible. Evidence that the factor of eye movement was controlled is based only on the reports of untrained Os. However, it will be remembered that they were told that if they made any eye movements from the fixation point they were to report them to *E*. All of the Os did this at one time or another which is some indication that they were cooperating. Also they were required to fixate for only short intervals at a time so it probably never became fatiguing. After each session the Os were asked if they had maintained their eye movement control and all reported they had, except the times indicated during the course of the experiment.

It is therefore concluded that the results of the study are evidence of the unity of the senses. This means only that the various sense modalities can be considered functionally as *not* completely independent systems. They might be considered as having 'membership character' in a larger, more inclusive whole in which they are related. However, in the adult human being, they represent relatively segregated systems. An event in one sense field is much more closely related in terms of mutual interaction to an event within the same sense field than it is to an event in another modality. Probably the two sense fields least differentiated from one another are taste and smell. Hornbostel (**6**) proposed the hypothesis that originally the various sense fields differentiated from a more primitive common sensorium. If this is true there would perhaps be some evidence for it in the phylogenetic development of the sense organs and their neural connections. It would follow also that children and lower animals should show more intersensory relations than adults because their sense fields are not yet so far differentiated.

As the various sense fields represent relatively segregated systems within a larger whole, intersensory effects will be limited. If a generalization could be hazarded at this stage of our knowledge, it would be that intersensory phenomena would appear under conditions in which the perception in the *influenced* modality is a labile, weakly structured configuration; and conversely, the configuration in the *modifying* modality is strongly structured, highly figured and yet possessing a strong 'tension,' *i.e.*, in some way it must be disequilibrated.

Zietz (12) discovered this to be true in connection with his investigations on intersensory relations.

The results of the present experiment seem clearly to fit this generalization. A flickering light is a very unstable yet highly 'figured' configuration and has strongly structured spatial relations. On the basis of purely auditory cues, sound localization is rather labile, on the average being accurate only to 5°. In this case the buzzer was of low intensity which enhanced the indefiniteness of its spatial location. In terms of Wheeler's (9) 'law of least action' perhaps the flickering light could be said to represent a point of 'low potential' in the perceptual field. Its intermittent flashing makes it a highly 'figural' region. It is the outstanding region of the visual field. If it represents, then, the point of 'low potential,' there will be a tendency, in a fluid field, for other events in the field to be influenced by it. Under these conditions the tendency toward closure, *i.e.*, the tendency toward stability of the perceptual field, results in the distortion of the phenomenal location of the sound.

In Condition III we have a different situation. In this case both lights are flickering; and seemingly, both should have equal effect. Instead we find the phenomenal sound source distorted toward the in-rhythm light. It might be suggested that these inter-modality relations organize under the same laws that perceptual organization follows within one modality, in which case the results of Condition III seem to represent a situation similar to that involved in Wertheimer's factor of 'common fate.' Wertheimer (as quoted from Ellis (8)) describes this factor as follows:

> A row of dots is presented:
>
> . . . .
> a b c d e f g h i j k l
>
> and then, without the subject's expecting it, but before his eyes, a sudden, slight shift upward is given, say to d, e, f, or to d, e, f, and j, k, l together. This shift is 'pro-structural' since it involves an entire group of naturally related dots. A shift upward of say c, d, e, or of c, d, e, and i, j, k would be 'contra-structural' because the common fate (*i.e.*, the shift) to which these dots are subjected does not conform with their natural grouping . . . .
>
> . . . *Sometimes a revolt against the originally dominant factor of proximity will occur and the shifted dots themselves thereupon constitute a new grouping whose common fate it has been to be shifted above the original row.*[3]

The experimental situation presented in this study is perhaps analogous to Wertheimer's dot patterns. Out of a dark background suddenly two lights appear. They stand out as a 'figure.' At the same time out of the 'ground' of stillness, a sound emerges. This concomitant pulsation of the two stimuli may be said to be their 'common fate' which is the condition for their 'closing' into a new pattern or whole.

[3] Italics are the author's.

In general, it can be concluded that the results add more evidence to the assumption that the different sense fields are basically and dynamically related. It has been shown that visual perceptions can affect apparent sound sources when past experiences and experientially established meaning factors are eliminated, or at least reduced to a minimum. The sense modalities can be represented as relatively segregated fields within a larger whole in which they are related.

Furthermore, the study presents an experimental technique which with certain modifications suggests a method of experimentally investigating a number of problems of intersensory relations. If the sound source were controlled more exactly, *e.g.*, by sending the sound to the O's ears through tubes, one of which can be lengthened or shortened (as was done by Goldstein and Rosenthal-Veit (**2**)), it might be possible to make exact psycho-physical measurements with this type of set-up. With this method of quantifying the strength of the closure tendency, it would be possible to make comparative studies to determine the differential effects under different field conditions.

Summary

1. The problem of this study was concerned with the influence of visual perceptions on sound localization.

2. The experimental set-up consisted of presenting two lights at spatially separated positions and also sounding a buzzer from various positions between the lights. Os made judgments of 'left,' 'right,' or 'middle' depending on the region from which they heard the sound. One of the lights flickered in the same rhythm as the buzzer sounded.

3. The results showed a tendency for the Os to skew their judgments of the sound's location toward the flickering light or the in-rhythm light as the case may be. The reliability of these results receives most of its weight from the consistency of this tendency. However, critical ratios calculated from the data showed at least half of the differences to be statistically reliable. The plausibility of such factors as conscious association, experientially established meanings, and eye movement were reduced as explanations of the results.

4. The results were interpreted as being a phenomenon of intersensory relations. Granting the functional unity of the senses the general conditions under which the phenomenon appears are a strongly structured, disequilibrated configuration in the *influencing* modality and a weakly structured, labile configuration in the *influenced* modality. It was suggested that some of the laws of perceptual organization worked out for events within a modality may apply in connection with these intersensory organizations.

(Manuscript received September 4, 1940)

References

1. Ewert, P. H., A study of the effect of inverted retinal stimulation upon spatially coördinated behavior, *Genet. Psychol. Mon.*, 1930, 7, 177–363.
2. Goldstein, K., and Rosenthal-Veit, Ueber Akustische Lokalisation und deren Beeinflussbarkeit durch andere Sinnesreize, *Psychol. Forsch.*, 1926, 8, 318–335.
3. Hartmann, G. W., II. Changes in visual acuity through simultaneous stimulation of other sense organs, *J. exp. Psychol.*, 1933, 16, 393–407.
4. ——, The facilitating effect of strong general illumination upon the discrimination of pitch and intensity differences, *J. exp. Psychol.*, 1934, 17, 813–822.
5. ——, *Gestalt Psychology*, The Ronald Press Co., New York, 1935.
6. Hornbostel, E. M. v., Die Einheit der Sinne, *Melos. Zeitschr. der Musik*, Vol. IV, 290–297. English translation by Elizabeth Koffka and Warren Vinton, The unity of the senses, *Psyche*, 1927, 7, 83–89.
7. Stratton, G. M., Vision without inversion of the retinal image, *Psychol. Rev.*, 1897, 4, 340–360.
8. Wertheimer, M., Laws of organization in perceptual forms, in Ellis's *Source Book of Gestalt Psychology*, 71–88.
9. Wheeler, R. H., *Science of Psychology*, T. Y. Crowell Co., New York, 1940 (Revised Edition).
10. Willey, E. F., Inglis, E., and Pearce, C. H., Reversal of auditory localization, *J. exp. Psychol.*, 1937, 21, 114–130.
11. Young, P. T., Auditory localization with acoustical transposition of the ears, *J. exp. Psychol.*, 1928, 11, 399–429.
12. Zietz, K., Gegenseitige Beeinflussung von Farb- und Tonerlebnissen: Studien ueber experimentell erzeugte Synäesthesie, *Zsch. f. Psychol.*, 1931, 121, 257–356.

THE SPECIFICITY OF THE EFFECT OF SHOCK FOR ERROR IN MAZE LEARNING WITH HUMAN SUBJECTS

BY JACK BERNARD AND RALPH W. GILBERT

The numerous investigations which have been made of the effect of mild electric shock administered for error in the maze situation have seemingly failed to demonstrate any specificity of action. Muenzinger (**7**) has stated at the conclusion of his report on a study of the effect of shock upon the learning of rats in a visual discrimination situation that ". . . the function of moderate electric shock . . . is general rather than specific, . . . it affects the total performance rather than the part responses it accompanies. In other words, its function is to make the animal respond more readily to the significant cues in the learning situation rather than that of inhibiting or facilitating the response which was shocked." A general effect of shock has been repeatedly demonstrated for human subjects, in the sense that such 'punishment' for error causes more rapid error elimination in a maze as a whole. Gilbert has demonstrated (**4, 5**) that to be effective shock need not be given immediately following any particular error. What has not yet been determined is whether the effect of shock can be specific as well as general, or is general only as Muenzinger argues.

If shock is given for every error everywhere in the maze, its effect might be wholly general. However, the results of previous experiments strongly suggest that there is also a specific effect. The evidence is qualitative rather than quantitative, being in the nature of the verbal reports of subjects. What is needed to show that shock given for a specific error has a direct effect upon the likelihood of repeating that error on later trials, is to give shock consistently for errors in some blind alleys, and to withold the shock consistently when errors are made in others, this with the same maze and the same subjects. In addition, we must not score repeated entrances into a blind alley as separate errors. If this were to be done, there would undoubtedly be a smaller total number of errors for shock alleys than for non-shock ones, but this difference in favor of the shock alleys might be due only to the relative absence of repeated entrances within single trials, and it could not be shown to be due in any part to a specific effect of shock *which carried over from one trial to another*. And unless a specific effect does carry over in this way, it is not an effect which is significant for *learning*. The only way of

eliminating the possibility refered to above is to disregard repeated entrances in scoring.

Subjects, Apparatus and Procedure

Subjects.—The subjects were fifty-two undergraduate students, twenty-six males and twenty-six females, at the Washington Square College of Arts and Science, New York University. They were divided into two groups, equated for sex, and designated Shock 'O' and Shock 'X.'

Apparatus.—The maze pattern employed was chosen with the aid of a small group of subjects who learned several mazes to completion. The maze finally adopted was that which took longest to learn (and that which therefore gave an opportunity of studying learning over a long period), and which showed the least differential difficulty of blind alley elimination. This was the Fox maze, the pattern of which is shown in Fig. 1. The actual size was 12.5 × 10 inches,

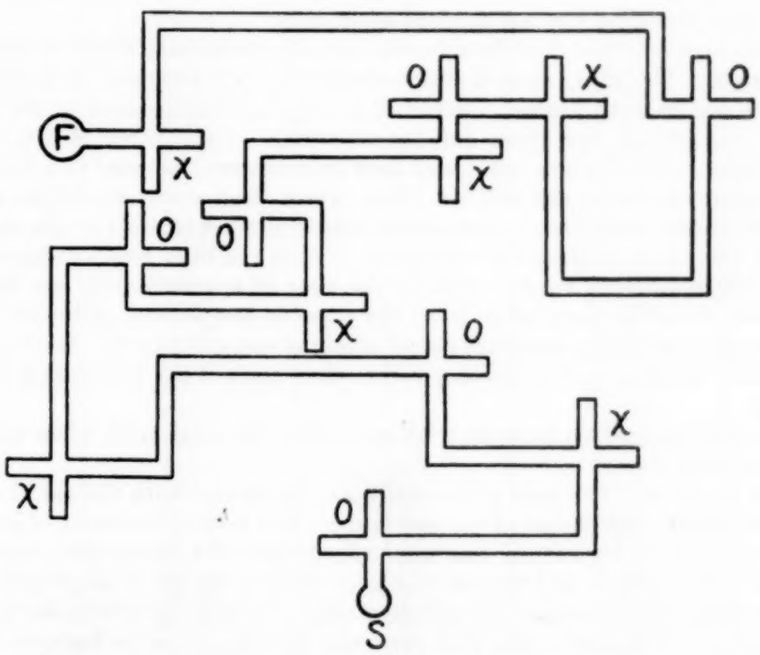


Fig. 1. Maze pattern employed in the present experiment.

the track was one fourth inch wide, and every blind alley was 3/4 inches long. The alleys were cut through a bakelite plate 1 cm. thick, and mounted on a glass base to reduce stylus friction. The blind alleys were arranged in pairs, and as shown in the diagram, alternate pairs of culs-de-sac were designated 'O' and 'X.'

The shock was administered to the subject's active hand by means of an electrical stylus and apparatus identical with that employed by Crafts and Gilbert (3), save for the following differences: 1. The Porter inductorium was placed in the room with the subject, since a silent shock was not imperative. 2. The experimenter controlled the shock by pressing a telegraph key, and counted errors by direct observation instead of employing an electric counter. Since it was desired to count initial blind alley entrances without counting repeated entrances, the counter could not be used.

A sample maze was shown to each subject when the problem was explained to him. Curtains concealed the maze and the experimenter from the subject.

Procedure.—The subjects of the Shock 'X' group received a shock each time the stylus point came in contact with the end of a cul-de-sac designated as 'X,' while the subjects of the Shock 'O' group received shock upon reaching the ends of the 'O' culs-de-sac. Thus when the results for these two groups were averaged as 'shock' and 'non-shock' scores respectively, the effects of any difference in intrinsic difficulty between 'X' and 'O' culs-de-sac were distributed

equitably between the 'shock' and 'non-shock' conditions. Only 'contact' errors were counted (errors involving the contact of the stylus tip with the end of the blind alley), and of these only the first entrance into a particular cul-de-sac during any one trial was recorded. Thus, by counting multiple errors as single errors (multiple errors involving repeated entrances into a blind alley immediately after having left it), the scores are rendered more significant as measures of the specific effect of the shock upon learning, as noted above. Although multiple errors were reduced to single errors in the record, shocks were given each time the stylus came in contact with the end of a shock alley, whether or not the entrances were multiple.

The amount of shock given each subject was the maximum that could be administered without producing effects obviously detrimental to the purpose of the experiment. This amount was determined for each subject by the method reported in an earlier paper (3).

In order to avoid having one error lead to several further errors in numbers out of proportion to its significance, retracing was prevented. The experimenter said 'Stop' each time it became apparent that the subject was retracing. Each subject was informed at the beginning of the experiment that he was to reverse the direction of his movement upon hearing the command 'Stop.' Attempted retracing was not counted as an error.

The subjects were informed that the criterion of complete learning would be three consecutive trials without error. However, since it was sufficient for our purposes to investigate a fairly large segment of the learning process, instead of having learning proceed to the satisfaction of this criterion, each subject was given only twenty trials. Only three of fifty-seven subjects satisfied the criterion within twenty trials, and their records were discarded to avoid the weighting of the average group scores in the last few trials, where their scores would have dropped out. Two other subjects were unable to complete the required number of trials within the time allowed (one hour), and their records were also discarded. Thus the final group consisted of fifty-two subjects, all of whom had run twenty trials in the time of approximately one hour. They all believed that they would be required to learn the maze to completion, and most of them manifested obvious surprise when the work was halted after the twentieth trial. Motivational orientation was therefore toward complete learning rather than toward the completion of the required number of trials.

No rest intervals were given between trials save after the tenth trial, when each subject was permitted two minutes rest.

Instructions to the subjects were given orally and illustrated with the aid of a 'multiple T' maze of a pattern totally unlike that of the test maze. The lack of similarity of pattern between the maze seen and that to be learned was emphasized when the instructions were given. The subjects were told that they would receive shocks in some of the blind alleys and not in others, but they were warned that entrances into *all* blind alleys, whether involving shock or not, would be counted as errors. They were told that their sole aim should be to learn to avoid entering blind alleys, and that all other factors, including time, were to be ignored. In this respect the instructions differed greatly from those usually given in stylus maze experiments, in which the subjects are given the impression that time is an important factor. The time required for each trial was recorded without the knowledge of the subject, and total time scores for each set of twenty trials were computed, but they are not included in this report since they seem to have little significance for this study, and do not appear to correlate significantly with any of the other variables.

Results

The criteria used for comparing the learning of shocked and unshocked alleys were: (1) The total number of errors made on the shocked and unshocked alleys. (2) The number of subjects making fewer errors in shocked alleys. (3) A measure of progress in learning, namely, the ratio of errors during the first 10 trials to the errors during the last 10 trials.

Table I presents ratios of errors on the first 10 trials to errors on the second 10 trials, for both shocked and unshocked alleys These ratios are given in the form of common fractions as well as in the

TABLE I

GIVING THE NO. OF ERRORS, RATIOS OF ERRORS MADE ON THE FIRST TEN TRIALS TO THE ERRORS MADE ON THE SECOND TEN TRIALS, FOR FIFTY-TWO SUBJECTS

Subj.	Shock		Non-shock	
	Errors 1st 10 trials / Errors 2nd 10 trials	Ratio	Errors 1st 10 trials / Errors 2nd 10 trials	Ratio
1	39/35	1.12	55/16	3.44
2	36/8	4.50	46/19	2.42
3	55/44	1.25	80/50	1.60
4	34/27	1.26	45/33	1.36
5	45/15	3.00	63/39	1.62
6	64/50	1.28	63/41	1.54
7	38/7	5.43	35/12	2.92
8	63/21	3.00	57/51	1.12
9	53/42	1.26	48/47	1.02
10	44/40	1.10	42/34	1.24
11	39/9	4.33	48/23	2.09
12	58/38	1.53	62/57	1.09
13	58/28	2.07	69/48	1.44
14	42/19	2.21	54/34	1.59
15	53/17	3.12	75/28	2.68
16	51/21	2.43	60/22	2.73
17	56/15	3.73	64/22	2.91
18	59/21	2.81	62/39	1.59
19	59/47	1.26	61/59	1.03
20	57/27	2.11	40/49	0.82
21	42/23	1.83	51/26	1.96
22	29/7	4.14	44/10	4.40
23	47/36	1.31	82/81	1.01
24	31/12	2.58	66/29	2.28
25	52/33	1.58	53/43	1.23
26	48/35	1.37	57/22	2.59
27	62/38	1.63	66/36	1.83
28	58/40	1.45	51/39	1.31
29	83/51	1.63	69/44	1.57
30	48/28	1.71	42/19	2.21
31	63/30	2.10	57/47	1.21
32	54/37	1.46	46/36	1.28
33	46/10	4.60	42/11	3.82
34	59/38	1.55	57/49	1.16
35	56/38	1.47	61/50	1.22
36	57/26	2.19	61/35	1.74
37	45/23	1.96	60/36	1.67
38	55/14	3.93	36/11	3.27
39	55/52	1.06	74/64	1.16
40	59/56	1.05	75/64	1.17
41	69/37	1.86	79/57	1.39
42	90/68	1.32	85/73	1.16
43	42/33	1.27	34/27	1.26
44	38/11	3.45	45/22	2.05
45	30/9	3.33	49/24	2.04
46	52/27	1.93	83/59	1.41
47	66/57	1.16	81/66	1.23
48	61/37	1.65	59/52	1.13
49	45/27	1.67	54/31	1.74
50	66/62	1.06	81/69	1.17
51	72/57	1.26	81/77	1.05
52	36/15	2.40	37/26	1.42

form of decimal fractions in order that the reader may be provided with the raw data. The total error score for any subject for the shock and non-shock alleys, respectively, can be obtained by adding the numerators and denominators of the common fractions given in the table under the headings Shock and Non-shock. Thus subject No. made 39 + 35 errors or 74 errors on the shock alleys, and 55 + 16 = 76 on the non-shock alleys. The total number of errors for all alleys is of course the total of the errors on the shock and non-shock alleys.

It will be noted that a ratio of more than 1.00 indicates fewer errors in the second ten trials than in the first ten, whereas a ratio below 1.00 indicates the reverse. A comparison of shock and non-shock ratios in any individual case will show whether that subject learned to avoid one type of error more rapidly than the other, and if so, how much.

In Table II will be found the mean ratios (errors on first ten trials to errors on second ten trials), and the mean error scores for shock and non-shock alleys, together with the standard deviations, differences and critical ratios.

TABLE II

	Shock Alleys		Non-shock Alleys		Difference	$\frac{\text{Diff.}}{\sigma_{\text{Diff.}}}$
	Mean	σ	Mean	σ		
Ratio 1st 10 Trials to 2nd 10 Trials.....	2.15	1.08	1.76	.79	.39	2.10
Errors.....	83	26.00	98	29.9	15.00	2.7

Small but fairly reliable differences are to be noted in favor of the learning of the shock alleys. Our measure of rate of progress in learning (the mean of the ratios described above) for the shock alleys exceeds that for the non-shock alleys by .39, the critical ratio being 2.1. This, as usually interpreted, indicates that the chances are about 98 in 100 in favor of a true difference greater than zero and in the direction indicated. The mean number of errors made on non-shock alleys exceeds the mean for shock alleys by 15, with a critical ratio of 2.7.

Thirty-eight of the 52 subjects made a larger number of errors on the non-shock alleys than on the shock alleys. For thirty-six of the 52 subjects the rate of learning was greater for the shocked than for the un-shocked alleys. The standard errors of these proportions are .062 and .064 respectively, with critical ratios of 3.72 and 2.96, making it seem improbable that repetition of the experiment under the same

conditions would yield either no significant differences or differences in the opposite direction.

The learning curves of Fig. 2 are simple 'error-trial' curves, plotted upon error scores as averaged by trials. They indicate that for the group as a whole learning as measured in terms of the elimination of errors was superior throughout the twenty trials for those alleys in which shocks were administered. As we might expect, if the shock had a specific effect, the curves diverge during the first few trials, and fail to cross at any point.

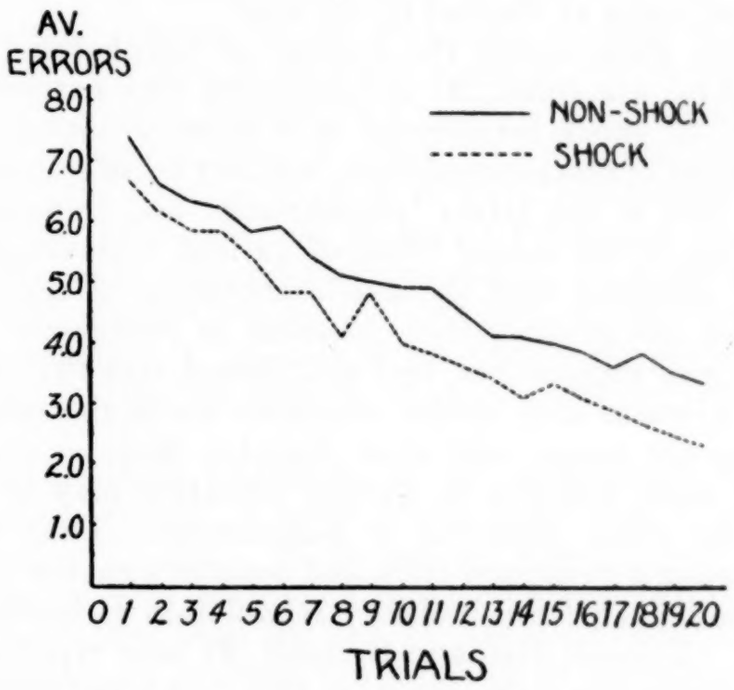


FIG. 2. Comparison of learning curves for shock and non-shock errors.

DISCUSSION

The data reported above would seem to indicate that, in addition to any *general* informational effect serving to prevent repeated entries in blind alleys on any one trial, and any general 'awakening' effect which shock may have, shocks administered for errors in maze learning also have what might be called a *marking effect.* This marking effect is specific, since to a certain extent it results in the singling out for earlier learning of those blind alleys in connection with which shocks are given.

Perhaps, from a broad point of view, the 'marking' effect is simply the awakening effect mentioned by Gilbert (4), occurring under somewhat altered conditions. In his 'non-informative' shock experiments, Gilbert applied shock while the subject was on the true path, or during the pause between trials, the number of shocks being

roughly proportional to the number of errors committed. Shock given in this way seemed often to reinstate the attentive attitude and thus facilitated learning. In the present experiment this reinstatement of an attentive attitude occurred each time the subject entered certain blind alleys. The shock called the subject's attention to the fact that he had made an error in this particular spot. It may have caused him to react more definitely and more vigorously to cues associated with the particular blind alley. In fact, the reports of subjects often indicated that several un-shocked errors might be made quite unconsciously, but a single shocked error was often remembered and remarked upon at the end of the trial.

We might, then, accept the concept of 'emphasizer' as used by Tolman and his associates (**8**) in connection with extraneous stimuli such as bell or shock introduced in a maze situation as constant accompaniment of certain responses, whether these responses be right or wrong. Use of the terms 'punishment' and 'reward' have impeded analysis of the actual effect of various extraneous stimuli on learning by carrying with them a connotation of the existence of stimuli which are of themselves pleasant or unpleasant. We often fail to take into account the well established fact that while in one situation a stimulus may arouse responses which the subject reports to be pleasantly toned, and that stimulus may be regarded as a reward, the same stimulus in another situation may be unpleasant and have the effect expected of punishment. Thus Gilbert and Crafts (**6**), using a combined click and buzz to signalize errors, found that subjects reported this auditory stimulation as having a punishment value. Tolman, Hall and Bretnall (**8**) have reported shock for right responses in the maze situation to act as a reward, while these same investigators and others (**1, 2, 3**) have found shock for wrong responses to have the effect usually expected of a form of punishment.

From the evidence available it seems reasonable to postulate that any well defined stimulus introduced consistently in conneection with either right or wrong responses will tend to favor their repetition if they are right or their elimination if they are wrong, provided that the stimulus is not of such type or strength as to introduce a distracting effect. In other words, any stimulus which is not highly distracting may act either as a 'punishment' or a 'reward,' depending upon whether it accompanies responses which are arbitrarily designated as right or wrong. This postulation is offered only for human learning, but quite probably it applies also to animal learning whereever a connection can be established either between the stimulation and blocking or the stimulus and the incentive.

Tolman and his associates have explained their 'emphasizing' effect in much the same manner that we have interpreted the 'mark-

ing' effect, but they have discussed this effect in connection with adient rather than avoidance reactions. Would not the same explanation apply to responses of both types? Whether we have adience or avoidance with respect to stimulus cues depends upon the situation in which the stimulus is applied rather than upon the stimulus itself, as suggested heretofore. With human subjects, verbal explanation of the significance of the shock (or other stimulus) immediately gives it meaning as reward or punishment. When explanation is not given to human subjects at the beginning of the experiment, or when animals are used as subjects, shock acts as a nocuous stimulus, and will almost invariably produce at first a withdrawal reaction. If this withdrawal reaction results in 'success' it will persist as a normal reaction to a nocuous stimulus. If it results in conflict, as when it is given in connection with a correct response, the shock may come to serve as a substitute stimulus for correct, goal approaching responses, and through a process of reconditioning, the normal avoidance reaction may be replaced by an adient reaction. No complete theoretical explanation in terms of conditioning will be attempted here, since in the case of the human subject the situation is complicated by the presence of implicit verbal or other symbolic responses.

In human subjects, knowledge of the significance of a 'punishment' stimulus appears to mean much more than direct affective reactions to the stimulus. In the present investigation most of the subjects reported that they tried to avoid shock mainly because it signified error rather than because it was disagreeable, in spite of the fact that they all reported the shock to be decidedly uncomfortable. Their reports also indicated that they tried to avoid un-shocked blind alleys as instructed, but that they tended to forget where these alleys were in the maze and that they often failed completely to perceive that they had entered a blind alley. On the other hand, the shock which was given in the 'shock' alleys announced the errors in a definite and clean-cut manner and enabled the subjects to mark the location more accurately.

Summary

In this experiment an attempt was made to determine whether, in a stylus maze situation in which moderate shock was given for entrance into some blind alleys but not for others, a different rate of learning and different error scores would obtain for shock alleys as opposed to non-shock alleys. Fifty-two subjects, twenty-six male and twenty-six female, ran twenty trials each in the Fox maze. Learning was in no case complete. Possible differences in difficulty of blind alley mastery were controlled by having the shock alleys for

half the group serve as the non-shock alleys for the remainder of the subjects.

The results indicate reliable differences in learning rates and simple error scores in favor of the shock alleys. Larger proportions of the group made higher learning rates and smaller error scores on the shock alleys, and the standard errors of these proportions suggest that further experimentation would yield differences in the same direction. This differential effect of shock for error is interpreted as showing that such shock has a specific effect upon the modification of responses to blind alleys in the stylus maze situation, in addition to the general effects upon learning which have been demonstrated by earlier investigations.

(Manuscript received August 28, 1940)

References

1. Bunch, M. E., Certain effects of electric shock in learning a stylus maze, *J. comp. Psychol.*, 1935, **20**, 211 ff.
2. Bunch, M. E., and McTeer, F. D., The influence of punishment during learning upon retroactive inhibition, *J. exper. Psychol.*, 1932, **15**, 473–495.
3. Crafts, L. W., and Gilbert, R. W., The effect of punishment during learning upon retention, *J. exper. Psychol.*, 1934, **17**, 73–84.
4. Gilbert, R. W., The effect of non-informative shock upon maze learning and retention with human subjects, *J. exper. Psychol.*, 1936, **19**, 456–466.
5. Gilbert, R. W., A further study of the effect of non-informative shock upon learning, *J. exper. Psychol.*, 1937, **20**, 396–407.
6. Gilbert, R. W., and Crafts, L. W., The effect of signal for error upon maze learning and retention, *J. exper. Psychol.*, 1935, **18**, 121–132.
7. Muenzinger, K. F., Motivation in learning. I. Electric shock for correct response in the visual discrimination habit, *J. comp. Psychol.*, 1934, **17**, 267–277.
8. Tolman, E. C., Hall, C. S., and Bretnall, E. P., A disproof of the law of effect and a substitution of the laws of emphasis, motivation and disruption, *J. exper. Psychol.*, 1932, **15**, 601–614.

FACTORS DETERMINING THE DIRECTION OF THE VISUAL AFTER-IMAGE DRIFT

BY THOMAS G. HERMANS

University of Washington

In classroom demonstrations of the visual after-image the question usually arises as to the cause of the apparent movement of the after-image. This movement, best described as drifting, seems quite involuntary and somewhat uncontrollable. It is a truism that the after-image is the result of retinal stimulation and a correlate of the persistent chemical effects in the retina continuing after the cessation of stimulation. It is also well established that the apparent movement of the after-image in space is a concomitant of eye movement. There is, however, no satisfactory explanation in the literature of the direction of this movement.

Rexroad [1] considers the drift a heliotropism. He reported that when the point of fixation during retinal fatigue was to one side of an illuminated figure the drifts were *always* in the direction of the greater illumination, and when imposed eye movements were in a direction counter to the expected heliotropic drifts no after-images were reported.[2] Gundlach and Walters [3] repeated Rexroad's experiment and found only 39 of the 65 reported drifts in the direction of the heliotropism Rexroad mentioned. They also found subjects were able to hold the after-image while movement of the eyes was opposite to the direction of the expected heliotropism.

Describing the drift tendency as a heliotropism is not particularly to the point inasmuch as the after-image may drift even though the retinal illumination was uniform or symmetrical relative to the point fixated during retinal fatigue. Further, the accurate use of the term heliotropism always implies either a positive or negative tendency. To use this term one would have to show that the drift was in one direction when a bright figure is used, while in the opposite direction when a dark figure is used. The use of this term is complicated further by the fact that it is the negative after-image that drifts. If it turned out to be a positive tropism relative to the light intensities used for retinal fatigue, it would be a negative tropism relative to the

[1] Rexroad, C. N., Eye-movements and visual after-images, *Amer. J. Psychol.*, 1928, 426–433.

[2] Following Washburn's suggestions in 'Movement and mental-imagery' in which she makes out a plausible case for the identity of movement and the memory image, Rexroad interprets this as proof of the identity of the after-image and eye movement. This error of juxtaposition was adequately refuted by Gundlach and Walters.

[3] R. H. Gundlach and V. W. Walters, Eye-movements and Visual after-images, *Amer. J. Psychol.*, 1931, 43, 288–289.

intensities in the after-image. Moreover, the term tropism is properly applied to the action of organisms that have no central nervous systems; such actions of organisms with a central nervous system might better be referred to as reflexes. The drift tendency observed by Rexroad, we shall refer to as the retinally determined drift, or the retinal tendency of the drift.

There is another factor which might plausibly determine the direction of the after-image drift, namely, the pattern of muscular tensions present during the retinal stimulation which is necessary for the production of the after-image. This is suggested by three different but related phenomena mentioned in the literature of muscle action. Troland [4] says, in a discussion of 'relative kinæsthetic sensation,'

> Allers and Saito find that the sensibility to differences in weight is increased by a single vigorous muscular effort, but is decreased by continued exertion. Borak notes that the sensibility is better when the lighter of two weights is lifted first than when the opposite procedure is employed. This may perhaps be attributed to adaptation of the tensive sense. Stand erect with the arms straight at the sides, facing along the length of a wall. Press the back of the right hand against the wall for a minute or two with strong exertion, using the muscles of the whole arm. Then step away from the wall. It will be found that if the arm is kept straight, it will rise from the side and come to equilibrium in an elevated position, the exact angle of which will be determined by the degree and duration of the previous exertion. . . . These phenomena can be explained on the assumption that the balance of innervation between the flexors and the extensors of the arm is determined by a corresponding equilibrium between the nerve currents from the proprioceptors of these muscles. The protracted pressure of the arm against the wall fatigues the receptors in the extensor system and allows those in the flexor system to recover unduly. Upon removal of the external restraint, the system automatically adjusts itself to a neurological balance, which demands increased tension in the tendons of the extensors.

This phenomenon might also be explained as E. B. Holt would explain it in terms of circular reflexes in muscles which serve to maintain their own contraction, rather than in Troland's terms of adaptation. Whichever explanation one uses, the example seems to have a close parallel in eye activity. The external muscles of the eyes may be under unequal strain while the retinas are being stimulated. The after effect is a tendency to continue the same movement. Troland [5] has made another observation that seems quite pertinent, but particularly so in the light of certain observations made by Hunter.[6] Troland states in a discussion of the after-image of visual movement:

> If, during the persistence of the motion after-image, an object is moved in opposition to it, a speed of the latter can be found at which the subjective impression is one of rest.

Hunter, in discussing the after effects of visual motion, noted that

[4] Troland, L. T., *The principles of psychophysiology*, D. Van Nostrand Co., Inc., 1930, Vol. II, pp. 343–344.

[5] Troland, L. T., *The principles of psychophysiology*, D. Van Nostrand Co., Inc., 1930, Vol. I, p. 387.

[6] Hunter, W. S., Retinal factors in visual after-movement, *Psychol. Rev.*, 1914, 31, 245–277.

there was a reduction of the apparent movement when the eyes during fixation of a point to the side of the moving visual field were in an extreme tension to the side opposite to the direction of the expected apparent after-movement. This effect was greater horizontally than vertically.

To determine whether the direction of the drift was due to the muscle tensions during the retinal stimulation or due to retinal factors the following experiment was devised. Several small groups of subjects, 46 in all, were recruited from elementary psychology classes. They were asked to report the initial direction of any apparent movement of the after-image that might ensue after each of a series of situations in which the eye position and the point of fixation were systematically varied. The objects used to produce the retinal fatigue were eight-inch colored disks placed on the wall in the ordinary classroom condition of illumination. A different color was used for each eye position, and during the retinal stimulation the subjects were told the approximate color of the negative after-image they might expect to see. They were not told the point of the experiment, and were left to assume that the seeing of the particular color of the after-image had something to do with the experiment. During the retinal fatigue they were also diverted from the possible aim of the experiment by the experimenter's comments concerning the phenomenon of sensory adaptation occurring as illustrated by the fading of the quality of the particular color used.

The sequence of situations of eye positions and fixation points was such as to check the two factors that might determine the direction of the drift. The subjects fixated either the center of the disk or a point to the left or to the right of the margin, and the eye position was either extreme right or extreme left in successive situations. In the table below the six variations of eye position and point of fixation are indicated. The letters *L* and *R* indicate the eye position, the *O* indicates the disk, and the period indicates the point of fixation relative to the disk. For example, *L·O* means that the subject's eyes were in the extreme left position of orientation, with the point of fixation to the left of the disk; *L*⊙ means that the eyes were extreme left but the point of fixation centered upon the disk. Of the six situations 1 and 2 were presented four times, and the other four situations were presented twice; this made a series of sixteen situations presented to all of the 46 subjects. Within the series left and right eye positions were alternated. The subjects were asked to write after each situation the direction of the initial drift movement that might occur, indicating up, down, left, right, or variants of these terms. In tabulating their reports any indication of right or left drift, whether 'down right' or 'up left' etc., was counted as a right or left drift; reports of 'up,' 'down,' 'no drift' and 'no after-image' were tabulated under the heading 'No.'

TABLE I

Situation during Retinal Stimulation	Reported Drifts			Total
	L	R	No	
1. *L*⊙	76	47	61	184
2. *R*⊙	24	108	52	184
3. *L·O*	16	49*	27	92
4. *R·O*	11	66*	15	92
5. *LO·*	62*	14	16	92
6. *RO·*	61*	16	15	92
Total.....	250	300	186	736

The number of reported drifts, left or right, in each situation representing drifts in the same direction as the eye position during retinal stimulation is underlined in the table. The underlined items

compared with the non-underlined left or right drifts would show to what extent the muscular tensions determine the drifts. The starred items represent the drifts reported that illustrate after-image drift as a retinally determined tendency. In situations 1 and 2 the retinal tendency would not be involved for the fixation was centered upon the disks. In situations 3 to 6 inclusive the retinal tendency might be involved, for the points of fixation were to left or right of the disk. Combining the underlined items of situations 1 and 2 we find 184 (50 percent of the 368) reports as compared with 71 (19.2 percent) in support of the theory that the muscular tensions during retinal stimulation, when fixation is in an extreme left or right position of orientation, establish a predisposition of the drift in the direction of that extreme orientation; 113 or 30.7 percent were of the no drift catagory. The starred items combined show 238 (64.6 percent of 368) reports as compared with 57 (15.5 percent) supporting the theory that the drift is retinally determined; 73 or 19.8 percent reports show no drift. Combining the underlined items of situations 3 to 6 we find 160 (43.4 percent) favoring the muscular theory as compared with 135 (36.6 percent) opposing the theory.

From the foregoing data it would appear that there are at least two factors determining the direction of the visual after-image drift. With an extreme orientation to left or right during retinal stimulation there is a tendency for the drift to occur in the direction of that orientation. When the point of fixation is to the right or left of the object used to produce the after-image the drift tends to occur in the direction of the object relative to the point of fixation. When the two factors oppose one another the retinally determined response tends to be dominant.

Two ways of testing the hypothesis that muscular tensions during retinal stimulation determine the direction of the drift of the after-image occurred to the author. (1) To fixate a point to the extreme right or left requires greater muscular strain than to fixate a point half way between front fixation and extreme right or left. If this is true, as subjective estimates would lead one to believe, one would expect a greater support of the muscle tensions hypothesis in extreme left or right eye positions than from half right or half left positions during the retinal stimulation. (2) One might also expect to find that the longer the duration of retinal stimulation in a given muscular position of orientation of the eyes the greater the tendency of the after-image to drift in that direction.

With additional groups of subjects in the ordinary classroom situation these hypotheses were tested. One-hundred twelve subjects in four elementary psychology classes were used. Two of the class groups, 57 students, were given one-half minute exposures to

produce the after-image, and two groups, 55 students, were given one minute exposures—this on the assumption that one minute of muscular adjustment might yield different results than one-half minute. The other conditions for all of the four class groups were the same.

The subjects were told that they were to report the intitial direction of any apparent movement of the after-image following retinal stimulation while the head had been held in certain positions to effect certain positions of the eyes within the head. They were asked to see first how far left and right they could turn their head, and shoulders if necessary, and still see binocularly the colored disk; then they were asked to estimate a position half way between the direct front or center orientation and extreme left or right. They were told that they would assume one of these five positions as directed in each of the periods of retinal stimulation. The five positions of center, half-right, half-left, extreme right, and extreme left, were run in series and the series of five repeated. Thus each subject reported the results of ten periods of retinal stimulation. As in the previous experiment the subjects were not told the intent of the experimenter, and further, to avoid the possible effect of verbally suggesting the reporting of a left or right drift, the subjects were asked to draw arrows indicating the direction of the initial movement that might occur. The reports were tabulated as before, counting as 'no' drift, up or down movements, no movement, or no after-image. To check against the author's own biases that might influence the interpretation of indicated directions, an assistant who knew nothing of the author's theories at stake made a similar but independent tabulation of the reports. The assistant's tabulation yielded essentially the same figures as the author's but favored the author's theory more than his own tabulation. The assistant's figures are given in the following table.

TABLE II

Duration of Retinal Stimulation	One-half Minute				One Minute			
Reported Drifts	L	R	No	Total	L	R	No	Total
Eye positions								
Center	23	60	31	114	20	56	34	110
Right one-half	19	*72*	23	114	24	*52*	34	110
Left one-half	*41*	42	31	114	*32*	40	38	110
Right extreme	13	*75*	26	114	22	*57*	31	110
Left extreme	*58*	32	24	114	*40*	31	30	110
Total	154	281	135	570	147	236	167	550

Before analyzing these data for the possibility of the above hypotheses being validated, let us note the preponderance of right drifts as compared with the left drifts—26 percent left, 46 percent right. Apparently in the experimental situations, the classroom, there was some asymmetry of conditions producing a greater tendency to right drifts. One factor was that the daylight illumination of the room came from windows on the subjects' left. As any left drift of the after-image would move toward an area of greater illumination, one would expect the after-image to be less readily detectable under such conditions. It is plausible that many of the reported 'No' drift category which also included 'no after-image' might have been left drifts had the conditions of illumination been uniform for the visual field. About this condition the experimenter chose to do nothing. Another factor that might have been productive of the higher percent of right drifts was the fact that the experimenter stood to the left of the colored disk during the retinal stimulation, and when removing the disk at the end of the period approached from the left and retreated again to the left. It seemed that we might be dealing with a 'negative experimenter-tropism.' To check this factor as a possible determiner of the asymmetry of the drift frequencies, the experiment was repeated with 102 additional subjects. The experimenter stood always on the right of the colored disk during the retinal stimulation period. The duration of the period was forty seconds for all subjects. The other conditions of the experiment were essentially the same as for the previous group, and again, an assistant who knew nothing of the theory involved tabulated the subjects' reports. The following table shows the results of her tabulation.

TABLE III

Reported Drifts	L	R	No	Total
Eye Positions				
Center	53	83	68	204
Right one-half	50	91	63	204
Left one-half	76	76	52	204
Right extreme	44	105	55	204
Left extreme	105	41	58	204
Total	328	396	296	1020

These data show some reduction in the preponderance of right drifts (38.8 percent right, 32.1 percent left) although essentially the same as, or better consistent with the previous data, as will be shown

in later statistical analysis. The author cherishes no belief that this shift in the proportion of right drifts demonstrates what was referred to as a 'negative experimenter-tropism.' Other undetermined factors such as a variation in the daylight illumination of the classroom might equally well have determined this shift.

There is another way of evaluating this problem of the preponderance of right drifts. Granted that the uncontrolled variables in the experimental situation resulted in too large a number of right drifts, such variables might nevertheless have occurred in a manner quite stable and constant while the other factors were systematically varied. If the uncontrolled variables are consistent in the results produced, the proportion of left drifts following a given eye position, to right drifts following the opposite or mirror eye position, should be in the ratio of 328 to 396. The paired items of Table III which are joined by straight lines should show this mirror relationship. We may use the Chi-Square Test to discover if the obtained frequencies of the paired items depart significantly from chance occurrence on the hypothesis that these data represent essentially a mirror relationship. The expected frequencies would be calculated by adding the mirrored frequencies of drifts and reapportioning 328/724ths of their sum to the left drift category and 396/724ths to the right drift category. The chi-square derived in this manner could be regarded as an index of reliability. This was done and it was found that from the data of Table III the chi-square was 7.00; with four degrees of freedom that these data allow, it means that there are about 15 chances in 100 that the hypothesis of a mirror relationship is tenable. Testing the same hypothesis on the data of Table II, a chi-square of 8.3 was found with the data obtained with the group given one minute of retinal stimulation, and a chi-square of 13.29 with the data of the one-half minute group. The hypothesis is even less tenable with these data than with the former. The conclusion to be drawn from this analysis is that the conditions of the experiment were such as to yield a disproportionately high frequency of right drifts, but further, this asymmetry is significantly inconsistent in its occurrence.

With this asymmetrical occurrence of drifts and the apparent lack of control of experimental conditions, one wonders if the data can have any meaning or significance for any theory of muscular factors determining the direction of the after-image drift. One method of determining the significance of these data is again by use of the chi-square test. If muscular factors influenced in no way the direction of the drift, the frequency of occurrence in the three categories should be much the same for all eye positions. On the hypothesis that these frequencies occur by chance, the expected frequencies of drift follow-

ing each eye position should be in the same proportion as the total frequencies of the three categories. With the data of Table III, for example, the expected frequency of left drifts for each of the five eye positions would be one-fifth of the total 328 left drifts obtained; the expected frequency of right drifts, one-fifth of 396; of no drifts one-fifth of 296. With such total frequencies as the basis for calculating the expected frequencies, chi-squares were calculated for the data from the three groups represented in Tables II and III. The data derived from the group using the one-half minute duration of retinal stimulation yielded a chi-square of 71.28; with the forty seconds group the chi-square was 71.24; and with the one minute group the chi-square was 31.31. With 8 degrees of freedom allowed for each chi-square, its size shows in each case the hypothesis of chance occurrence of drifts to be distinctly untenable; with the least significant of these (31.31) there is not even 1 chance in 100 that it represents a chance occurrence of drift frequencies.

In calculating the chi-squares for the data of each group, it is very significant, in the light of the theory being tested, that the most marked differences between the observed frequencies and expected frequencies on the hypothesis of chance, occur in the category of left and right drifts when the eyes during retinal stimulation were in a position of extreme right or extreme left. The next in magnitude of differences occur in right and left drifts when the eyes were in a half right or half left position. The frequencies of the no drift category differ but little from the expected frequencies on the hypothesis of chance. In other words the very significant size of the chi-squares obtained is determined predominantly under those conditions which test the theory as to the muscular tensions during retinal stimulation being productive of the direction of the drift of the after-image.

Another way of evaluating the data to determine if the theory of muscular factors determining the direction of the drift is tenable, is to compare the number of reported drifts in support of the theory with the reported drifts occurring contrary to the theory. Table IV summarizes the data of Tables II and III. The column headed with a plus lists the drift frequencies in support of the theory; these are the items underlined in the previous tables. The column headed with a minus sign lists the frequencies of drifts occurring contrary to the theory. The critical ratios cited in the table pertain to differences in percent of drifts occurring under the different conditions of the experiment. Those at the bottom of the columns are the critical ratios pertaining to the differences between percentages of drift frequencies following one-half left or right orientation during retinal stimulation, and extreme right or left orientation. The

TABLE IV

Eye Position	Group of 57 Subjects. Thirty Seconds of Retinal Stimulation				Group of 102 Subjects. Forty Seconds of Retinal Stimualtion				Group of 55 Subjects. Sixty Seconds of Retinal Stimulation			
	+	−	No	Total	+	−	No	Total	+	−	No	Total
One-half of extreme right or left	113	61	54	228	167	126	115	408	84	64	72	220
				C.R.				C.R.				
Percent	49.5*	26.7	23.8	2.04	40.9*	30.8	28.1	0.7	38.1*	29.0	32.7	
Extreme right or left	133	45	50	228	210	85	113	408	106	53	61	220
				C.R.				C.R.				
Percent	58.3*	19.7	21.9	1.68	51.4*	20.8	27.7	0.9	48.1*	24.1	27.7	
Critical ratios	1.91	1.84			3.08	3.33			2.14	1.16		

critical ratios cited within the horizontal rows pertain to the difference between the starred items found on either side of each critical ratio.

We return now to the theories that we set out to evaluate. The first was the theory that, if muscle tension to left or right during retinal stimulation gave rise to a drift in the same direction of that muscle tension, one should expect with extreme orientation to right or left a higher frequency of drifts consistent with the theory than when orientation is only half way between front orientation and extreme right or left. The theory seems to be amply supported in the data of Table IV. With the data from each of these three groups differentiated on the duration of retinal stimulation we find the percent of drifts in support of the theory significantly higher when the extreme orientation left or right was used than when the orientation was half left or right. The smallest of these critical ratios is 1.91, which means that there are 97 chances in 100 that this is a significant difference, while the other critical ratios are even more significant. It is also highly important to note the percent of drifts in non-support of the theory reported by each group is significantly lower under conditions of extreme orientation left or right than under conditions of only half left or right orientation. The least significant of these critical ratios is 1.16, which means that there are 87 chances in 100 that this is a significant difference. If we combine the data from the three groups, disregarding for the moment differences in duration of retinal stimulation, we find very significant differences between the one-half left or right orientation and extreme left or right orientation. In support of the theory there were 42.52 percent for half orientation and 52.45 percent under conditions of extreme orientation; the ratio of this difference to the sigma of the difference is 4.39. The drifts reported in non-support of the theory showed 29.32 percent under conditions

of half orientation and 21.37 percent under conditions of extreme orientation; this difference in percent yields a critical ratio of 3.8. Such critical ratios show beyond a doubt the theory to be tenable; the farther the orientation left or right during retinal stimulation the greater the tendency of the drift to occur in that direction.

The second theory that we shall attempt to evaluate from the data of this experiment is that the longer the duration of retinal stimulation in a given muscular position of orientation the greater the tendency of the after-image drift to occur in that direction. Examination of the data of Table IV reveals quite the contrary to be the fact. Within the limits of time used, 30, 40, and 60 seconds, there was a very consistent decrease of the tendency of the drift to occur in the direction of the predisposing muscle tensions, with an increase in duration of retinal stimulation. Superficial scanning of Table IV shows this tendency to be quite consistent not only in the decreasing percent in support of the first theory with increase in time, but also increase in percent of drifts in non-support and no-drift category with increase in time. The four critical ratios cited in the table are, however, not particularly significant. Combining the data of Table IV vertically, in other words disregarding the distinction in degree of orientation left or right, we have a far more interesting comparison of this tendency. Table V shows this comparison.

TABLE V

	Thirty Seconds Group				Forty Seconds Group				Sixty Seconds Group			
	+	−	No	Total	+	−	No	Total	+	−	No	Total
Numbers	246	106	104	456	377	211	228	816	190	117	133	440
Percent of Total	A 53.9	A′ 23.2	A″ 22.8		B 46.2	B′ 25.8	B″ 27.9		C 43.2	C′ 26.5	C″ 30.2	
					AB 2.65	A′B′ 1.04	A″B″ 2.04		BC 1.03	B′C′ 0.26	B″C″ 0.85	
	Critical ratio of difference in percent to the sigma of the difference.								AC 3.24	A′C′ 1.17	A″C″ 2.55	

The significance of the difference in percent indicated in this table by the sizes of the critical ratios is not extremely great in most cases. It is, however, of considerable significance to note that every one of these nine critical ratios pertains to a difference in percent which is consistent with the general picture. That picture is one of a decreasing effectiveness of muscle tensions in producing drifts in directions consistent with such tensions, with an increase in duration of the retinal stimulation.

This tendency at first appeared to the author to be consistent with the retinally determined tendency of the drift observed when the point of fixation during retinal stimulation was to one side of the object. The increase in duration of retinal stimulation would increase the vividness of the after-image, and this would render the retinal factor more prominent, but since the fixation was centered the increased prominence of the retinal factor should be to decrease the tendency to drift in any direction. If this were true, one would expect to find with increase in duration of retinal stimulation an increase frequency of retinally determined drifts following fixation of points to one side of the object.

The relation between the duration of retinal stimulation and the frequency of retinally determined drifts was tested with an additional group of 55 subjects. Each subject observed eight after-images. Two observations were made with each of 15, 30, 45, and 60 seconds of retinal stimulation. For each time used the subject fixated a point to the left and then to the right of the colored disk. The other conditions of the experiment were essentially the same as in the previous experiments. Of the reported drifts 72 percent illustrated the retinal tendency and 11 percent were opposite in direction to the retinally determined tendency. The Chi-Square Test was used to determine if the obtained frequencies of drifts deviated significantly from chance. With six degrees of freedom that the data allowed, the chi-squared of 11.64 obtained could be considered of border-line significance on its face value. However, nearly half the size of this chi-squared is determined by many reports of no after-image resulting from only 15 seconds of retinal stimulation. The remaining data would not show a significant deviation from chance occurrence of drifts. The author pursued no further this phase of the problem, although it seemed plausible that one might find significant differences of drift tendency with differences in duration of retinal stimulation had other amounts of time been used.

Summary and Conclusion

The foregoing experiments demonstrate that there are at least two factors involved in the determination of the direction of the visual after-image drift. One is the point of fixation relative to the object used to produce the retinal effects necessary for the production of the after-image; or we might say, the position of the retinal effects relative to the fovea which we have called the retinal factor. The other factor is the pattern of muscular stresses during retinal stimulation which leave their after effects persistent during the experiencing of the after-image; this we have referred to as the muscular factor.

The retinally determined drift is demonstrated when the point of fixation is to one side of the object fixated, and the drift ensuing tends to be in the direction of the object relative to the point of fixation; this drift is dependent upon the locus of retinal stimulation and within the limits of the durations of retinal stimulation used is independent of the duration, and also independent of the muscular factors. The muscularly determined drift is demonstrated when the eyes during retinal stimulation are in a position of left or right orientation with fixation centered upon the object; this drift tends to be in the direction of the preceding muscle tensions. The greater the degree of this angular orientation the greater the tendency to this drift, but the longer the duration of the muscular strain, within the range of times used, the less the tendency of the drift to occur in the direction of such muscle strain.

(Manuscript received August 24, 1940)